Journal of Experimental Psychology

UNIVERSITY OF MICHIGAN
JUN 30 1955
SOCIAL SCIENCE LIBRARY

ARTHUR W. MELTON, Editor
AIR FORCE PERSONNEL AND TRAINING RESEARCH CENTER
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

CONSULTING EDITORS

Judson S. Brown, *State University of Iowa*
Cletus J. Burke, *Indiana University*
Robert M. Gagné, *AF Personnel and Training Research Center*
W. R. Garner, *The Johns Hopkins University*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
David A. Grant, *University of Wisconsin*
Lloyd G. Humphreys, *AF Personnel and Training Research Center*
Arthur L. Irion, *Tulane University*
Donald B. Lindsley, *University of California, Los Angeles*
Neal E. Miller, *Yale University*
Kenneth W. Spence, *State University of Iowa*
Benton J. Underwood, *Northwestern University*
Delos D. Wickens, *Ohio State University*

Lorraine Bouthilet, Managing Editor

B
1
.J8

CONTENTS

The Identification of Redundant Stimulus Patterns: P. D. BRICKER 73
Relation of Stimulus and Response Amplitude to Tracking Performance: B. O. HARTMAN AND P. M. FITTS 82
Compound Trial-and-Error Learning as a Function of Response Availability (N_R): C. E. NOBLE 93
A Test of a Model for Multiple-Choice Behavior: M. H. DETAMBEL 97
Stimulus Generalization of an Instrumental Response as a Function of the Number of Reinforced Trials: G. MARGOLIUS 105
The Effect of Learning Verbal Labels for Stimuli on Their Later Discrimination: J. S. ROBINSON 112
Adaptation to Time in a Sensorimotor Skill: R. CONRAD 115
Effects of Prolonged Exposure to Low Temperature on Visual-Motor Performance: W. H. TEICHNER AND J. L. KOBRICK 122
Rotary Pursuit Performance as Related to Sex and Age of Pre-Adult Subjects: R. B. AMMONS, S. I. ALPRIN, AND C. H. AMMONS 127
An Analysis of the "Vertical-Horizontal Illusion": T. M. KÜNNAPAS 134
The Variation in Width and Position of Mach Bands as a Function of Luminance: C. MCCOLLOUGH 141

American Psychological Association

Vol. 49 No. 2 February 1955

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $7.50, or $15.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This JOURNAL publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology. Normally, articles of a length exceeding 15 printed pages cannot be accepted. However, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) must be presented in a single article, rather than in a series of articles.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1955 the publication cost is $20 per printed page. Authors of priority publications receive no free offprints.

Address all articles submitted for regular or priority publication to the editor, Arthur W. Melton, Air Force Personnel and Training Research Center, Box 1557, Lackland Air Force Base, San Antonio, Texas until March 19, 1955. From March 19 to June 1 address manuscripts to the editor at Department of Psychology, Stanford University, Stanford, California. Manuscripts sent after June 1 should be addressed to San Antonio again. All manuscripts should be submitted in duplicate. Original figures are prepared for publication; duplicate figures may be photographic or pencil-drawn copies. Authors are cautioned to retain a copy of the manuscript to guard against loss in the mail.

To save printing costs supplementary material to articles in this Journal are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

Manuscripts submitted to the JOURNAL *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc. as described in the Publication Manual of the American Psychological Association (*Psychol. Bull.*, 1952, 49 [4, Suppl.], 389–449). In particular, the organization of the manuscript should be text, references, footnotes, tables (one to a page), figure titles (typed on separate page), and figures. All typed material *must* be double-spaced with ample margins. The term "subject," with its various forms, is abbreviated *S*, *Ss*, *S*'s, and *Ss*'; analogous abbreviations are used for "experimenter" (*E*) and "observer" (*O*). When in doubt about the practices of the JOURNAL, authors should examine a recent issue.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879. Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 25th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Copyright, 1955, by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 49, NO. 2 FEBRUARY, 1955

THE IDENTIFICATION OF REDUNDANT STIMULUS PATTERNS[1]

PETER D. BRICKER[2]

The Johns Hopkins University

When an observer can consistently or reliably make the correct response to one of several possible stimulus patterns, he is said to be able to identify the stimulus. Previous experiments have measured the rate of acquisition of this identifying response (**1**) and the reaction time required for it (**2, 3**). This experiment uses both of these measures as indices of the difficulty of identification.

Several stimulus factors have been shown to affect these measures of the identifying response. This experiment is concerned with the effect of having more than the necessary number of cues to the identity of each stimulus, when all the cues are of the same physical nature. The number of formal cues is easily varied when only the eight stimuli shown in Fig. 1 are used. A given stimulus in this list appears to *S* as five lighted glow tubes (represented by a given group of five circles within a rectangle) on an array of two rows of five tubes each. But from *E*'s point of view, a stimulus may be thought of as composed of five subareas, each of which is a vertical column of two possible lights. A given subarea can be in one of two states—either the top or bottom tube is lighted. Because there are just these two possibilities for any one unit, the units are referred to as binary elements. The state or value of each binary element is one cue to the identity of the stimulus.

When there are eight possible stimuli, only three such binary cues are required to identify a stimulus. In other words, the two lights at the extreme left can be eliminated from all eight stimuli in Fig. 1 without destroying the uniqueness of each stimulus. The basic three-element stimuli can be made to contain surplus cues in two ways. First, the number of binary elements can simply be increased, while the number of stimuli and the number of responses is held at eight. This kind of variation, which is called the *stimulus redundancy* variable, produces the four- and five-element stimuli shown in Fig. 1. It makes the stimulus patterns more complex, but also adds alternative ways of identifying each stimulus.

[1] Based upon a dissertation submitted to the Faculty of Philosophy, The Johns Hopkins University in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The author wishes to thank Drs. H. W. Hake and W. R. Garner for their advice and assistance in carrying out this project.

[2] Now at Control Systems Laboratory, University of Illinois.

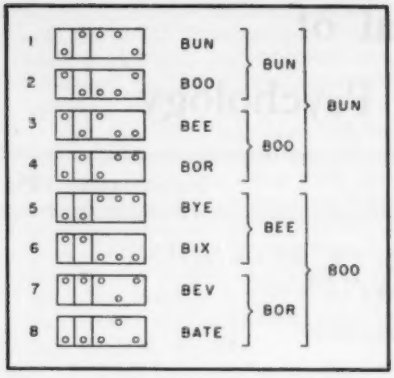

FIG. 1. The three lists of stimuli and the three lists of responses used in the experiment. An open circle represents a lighted glow tube, and an open space inside the rectangle and immediately above or below a circle represents an unlighted glow tube. In any one list of stimuli, all eight stimuli consist of the same number of lighted tubes. The appearance of a particular stimulus in a given list is indicated by the group of 3, 4, or 5 circles, beginning on the right, enclosed by that particular horizontal rectangle. The responses in the eight-response list were associated with the stimuli directly opposite. Those in the four- and two-response lists applied to those stimuli included in the brackets.

Second, the number of responses available to *S* can be reduced, while the number of stimuli and the number of elements are held constant. With four or two responses, associated with stimuli as indicated by the brackets in Fig. 1, *S* need only identify each stimulus as belonging to a class of two or four, respectively. This is the *response uncertainty* variable.

If one element is deleted from a nonredundant stimulus, the remaining cues are insufficient for identification. But redundant stimuli can still be identified under these circumstances, so the occasional deletion of one element provides a check on this predicted advantage of redundant stimuli. This is the *deletion* variable.

The experiment consisted of introducing these three variables into the task of learning to identify eight different stimuli. Its primary purpose was to determine whether alternative-cue redundancy facilitates the identification of stimulus patterns.

METHOD

Subjects.—Ninety male undergraduates at The Johns Hopkins University served as *S*s. They were all serving subject duty required of introductory psychology students. Each was assigned to a condition according to a predetermined random order.

Apparatus.—The stimuli appeared for .7 sec. on two rows of five neon glow tubes each. The sequences in which the stimuli appeared were punched in code on a teletype tape. This tape was read by an electronic sequence programmer which presented all of the trials automatically at 12-sec. intervals. The clock of an electronic voice reaction timer was started at the instant a stimulus appeared, and was stopped at the initiation of *S*'s vocal response.

Stimuli and responses.—The three lists of stimuli which correspond to the three conditions of the stimulus redundancy variable are shown in Fig. 1, where an open circle represents a lighted glow tube. The stimuli used in the condition of zero redundancy, each of which is composed of three elements, are shown in the three right-hand columns. Note that this list includes all possible arrangements of three binary elements. The four-element stimuli are shown in the four right-hand columns, and the five-element stimuli cover all five columns. In terms of the number of stimuli which could have been generated with four and five elements, these lists are 50% and 75% redundant, respectively. Stimuli in the redundant lists can be differentiated from the others in the list by the correct perception of *any* three elements. For this reason, the addition of redundant elements is considered to have the effect of providing alternative ways of identifying each stimulus.

Incomplete stimuli were produced by deleting one element. All of the stimuli were complete for half the *S*s (Cond. C), and 10% of the stimuli were incomplete for the other half (Cond. 10%-I).

The responses were all monosyllables beginning with the letter B. There were three lists of responses, composed of 8, 4, and 2 responses, respectively, which represented the three conditions of response uncertainty. When *S* has no knowledge of the stimulus-response relationships or of the structure of the stimulus series, 2, 4, and 8 response categories correspond to 1, 2, and 3 bits of uncertainty, respectively. The three lists of responses, together with the

stimuli for which they were the correct names, are shown in Fig. 1. Every *S* was provided with one of these three lists of responses and was required to associate each response in the list with one or more stimuli. The assignment of groups of stimuli to response categories for the two- and four-response lists was done in a way which maintained the difficulty of identification, on the basis of formal cues, at a maximum. In other words, there was an attempt to disguise the fact that less information was required for the identification of each stimulus as response uncertainty was reduced.

Procedure.—The three conditions of stimulus redundancy, the three conditions of response uncertainty, and the two conditions of deletion were factorially arranged. This provided 18 experimental conditions, with five *S*s in each. The condition to which *S* was assigned determined which list of responses he was given, which tape was used to program his stimuli, and to some extent, the instructions he received.

Before the experimental session began, each *S* was given a list of the responses he was to use and told what the general nature of the stimuli would be. He was told to respond as quickly as possible after each stimulus with one of the names on his list, even if he had to guess, and that he would be corrected if he were wrong. The *S*s in Cond. 10%-I were told that incomplete stimuli would occur. Then the apparatus automatically presented 200 stimuli. After each stimulus, *E* recorded *S*'s response and his reaction time, and pronounced the correct name of the stimulus.

Each successive group of eight stimuli included all eight possible stimuli, but the order within groups was random. Thus, differences between the stimuli within a list are equalized for the successive blocks of eight trials. The incomplete stimuli were distributed throughout the series so that one occurred in each ten successive presentations. The stimulus affected and the element deleted were randomly determined.

Results

Two main measures were used in this analysis. The first of these measures, which is referred to hereafter as the *per cent correct*, is the corrected percentage of a certain number of trials on which *S* made the correct response. The per cent correct was calculated from a number-correct score which had been corrected for chance. This corrected accuracy score was found by the formula:

$$R_c = R - W(p_r/p_w),$$

where R_c is the corrected number right, R is the raw number right, W is the number wrong, p_r is the probability that a correct response will occur by chance, and p_w is the probability that a wrong response will occur by chance. The correction is made separately for each condition of response uncertainty because the probability that a correct response will occur by chance alone increases as response uncertainty decreases.

TABLE 1

Analysis of Variance of Reaction Time and Error Scores

Source	*df*	Reaction Time			Total Errors		
		Mean Square	*F*	*p*	Mean Square	*F*	*p*
Stimulus redundancy	2	3,926.35	11.64	<.001	6,371.08	10.97	<.001
Response uncertainty	2	16,507.88	48.93	<.001	2,239.85	3.86	<.05
Deletion	1	1,777.78	5.27	<.05	4,354.18	7.49	<.01
Stimulus redundancy × Response uncertainty	4	339.66	1.01	—	1,327.81	2.29	.05<*p*
Stimulus redundancy × Deletion	2	1,092.08	3.24	<.05	307.46	—	
Response uncertainty × Deletion	2	80.75	—		1,330.98	2.29	.05<*p*
Stimulus redundancy × Response uncertainty × Deletion	4	266.19	—		139.07	—	
*S*s within groups (error)	72	337.39			580.99		

The other important measure is the mean voice reaction time for the last 50 trials. It is interpreted here as a fairly stable and asymptotic measure of reaction time because no statistically significant trend appears in reaction time over these trials for any group. Both of these measures were treated by the analysis of variance, the results of which are shown in Table 1. Note that the accuracy measure actually used was the corrected number wrong, which is the exact complement of R_c in the above expression. The reader should refer to this table for details of the statements about significance made in the text.

Stimulus redundancy and the course of learning.—The learning functions for each of the three conditions of stimulus redundancy are presented in Fig. 2 and 3, where the per cent correct is shown for each successive block of eight trials. The data for 15 *Ss*, five for each number of response categories, are plotted at each point. The number correct in a given block was corrected for chance by the appropriate formula for each number of response categories separately, and these corrected scores were then averaged to determine per cent correct for all 15 *Ss*. The curves

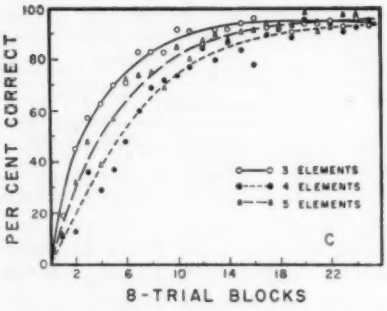

FIG. 2. Per cent correct responses in each successive block of eight trials, Cond. C only. Each point is an average over all conditions of response uncertainty.

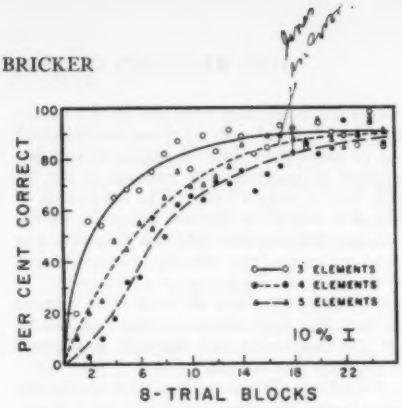

FIG. 3. Per cent correct responses in each successive block of eight trials, Cond. 10%-I only. Each point is an average over all conditions of response uncertainty.

differ chiefly in the rate at which they approach the asymptote, and all three stimulus redundancy groups within either Cond. C or Cond. 10%-I reach about the same final level of accuracy. The per cent correct in all 200 trials, therefore, is related directly to the rate of learning under each condition. The number correct, from which per cent correct is derived, is amenable to conventional statistical analysis.

The total per cent correct for each condition of stimulus redundancy is shown in Table 2. An analysis of variance shows the over-all differences among these three conditions to be significant, and to account for more variance than any other real source. Stimulus redundancy did not interact significantly with response uncertainty, so each value in Table 2 and

TABLE 2

PER CENT CORRECT IN 200 TRIALS AVERAGED FOR ALL CONDITIONS OF RESPONSE UNCERTAINTY

Deletion Condition	Number of Elements		
	3	4	5
C	82.5	71.0	77.9
10%-I	79.2	61.5	69.8

each point in Fig. 2 and 3 for a given number of elements is an average over the three conditions of response uncertainty. The curves and over-all per cent correct scores for Cond. C and 10%-I are shown separately, even though redundancy did not interact with deletion, because deletion had a significant effect.

These data show that *Ss* with nonredundant three-element stimuli learn faster than those with redundant stimuli. There is no evidence whatsoever that stimulus redundancy accelerated the learning process, although formally it could have provided additional cues to the identity of each stimulus. In fact, the relative standings of the three stimulus redundancy conditions on the per cent accuracy measure cannot be accounted for when the stimuli are specified in terms of redundancy. The stimuli may be specified also, however, in terms of the patterns formed by connecting the lighted glow tubes with imaginary straight lines. Obviously, the kinds of patterns formed depend on the number of elements in the stimulus. The assumption that *Ss* did organize the stimulus elements in this way permits some tentative predictions about the relative difficulty of identifying the stimuli in the three conditions. These predictions can be made directly from Fig. 1.

TABLE 3

NUMBER OF CORRECT RESPONSES TO EACH STIMULUS DURING THE FIRST 80 TRIALS FOR *Ss* USING EIGHT RESPONSE CATEGORIES ONLY

Number of Elements	Stimulus							
	1	2	3	4	5	6	7	8
3	63	39	42	53	77	71	70	74
4	63	37	33	30	36	40	55	52
5	64	49	31	57	54	46	48	51

The fifth and sixth stimuli in Fig. 1 form straight horizontal lines in the three-element condition, but this simple pattern is broken up when elements are added. The first prediction, based on the assumption that the more simple patterns are easier to learn, is that Stimuli 5 and 6 will be the easiest to learn when there are three elements, but no easier than the others when extra elements are added. The data in Table 3 essentially confirm this prediction.[3] This table shows the number of correct responses to each stimulus over the first 80 trials, for *Ss* with eight response categories, and for each stimulus redundancy condition separately. In the three-element condition, accuracy reaches a considerably higher level during these early trials for Stimuli 5 and 6 than for the first four stimuli, and Stimuli 5 and 6 average slightly higher than Stimuli 7 and 8. The superiority of the accuracy for Stimuli 5 and 6 is not as marked when there are four or five elements, however. Table 3 also shows that Stimuli 7 and 8 are easiest to learn under the three-element condition, where they form simple wedge-shaped patterns. The loss of simple figures, then, may be one reason why learning is slower with more than three elements.

Predictions can also be made about which stimuli will be confused with each other most frequently under each condition. It is reasonable to assume that stimuli in which the elements bear the same relation to each other, and which simply represent different orientations in space of the same figure, will be confused frequently.

[3] No attempt is made in this report to treat data of the type shown in Tables 3 and 4 comprehensively, because they are based on small numbers of cases. These samples are presented merely because they suggest explanations for the relative difficulty of identifying stimuli composed of different numbers of elements.

TABLE 4

THE FREQUENCY OF OCCURRENCE OF EACH RESPONSE TO STIMULUS 3 DURING THE FIRST 80 TRIALS FOR *Ss* USING EIGHT RESPONSE CATEGORIES ONLY

Number of Elements	Response							
	Bun	Boo	Bee	Bor	Bye	Bix	Bev	Bate
3	16*	9*	42	9*	3	6	11	4
4	4	6	33	13*	11	4	13*	16*
5	6	13	31	22*	7	8	7	6

Note.—Stimulus 3 was presented a total of 100 times to these *Ss* in each condition of stimulus redundancy during the first 80 trials.

* For a given condition of stimulus redundancy, the responses marked with an asterisk in that row correspond to those stimuli which are geometrically related to Stimulus 3 in the manner discussed in the text.

Figure 1 shows that the average stimulus in a given list is related in this way to the greatest number of other stimuli in the four-element condition, and to the least number in the five-element condition. An example of the obtained confusion patterns is shown in Table 4.[4] This shows the responses made to Stimulus 3 out of 100 presentations of that stimulus to *Ss* using eight response categories under each stimulus redundancy condition during the first 80 trials. The responses which would be expected to occur frequently under each condition, if confusions among geometrically related stimuli were important, are marked with an asterisk. Table 4 indicates that these confusions do occur frequently, and it has already been pointed out that the four-element condition affords the greatest number of opportunities for such confusions. There is good reason to believe, therefore, that the number of close relationships among stimulus patterns is another important factor determining the relative rates of learning under the three conditions of the stimulus redundancy variable.

The above two factors were illustrated with data from the group with eight response categories, but they would be expected to operate in a modified form with four response categories. The data for *Ss* with four response categories do in fact show the expected effects, and the factors may also be presumed to be operating when there are two response categories.

Stimulus redundancy and reaction time.—Mean reaction times for the last 50 trials, for each condition of stimulus redundancy and for Cond. C and 10%-I separately, are shown in Table 5. The over-all trend indicates the same order of difficulty of discrimination as shown by the per cent accuracy scores, and the same factors may account for it. The interaction of stimulus redundancy and the deletion variable approaches statistical significance, however, so it should not be overlooked. The difference between Cond. C and 10%-I is

[4] See footnote 3.

TABLE 5

MEAN REACTION TIME IN SEC./100 OVER THE LAST 50 TRIALS AVERAGED FOR ALL CONDITIONS OF RESPONSE UNCERTAINTY

Deletion Condition	Number of Elements		
	3	4	5
C	76.7	98.5	95.5
10%-I	91.6	115.3	90.5

significant for three and four elements ($df = 1$ and 24, $F = 5.91$ for three elements, $F = 4.54$ for four elements, $p < .05$ in both cases), but the apparent reversal at five elements is not significant ($F < 1.0$). The data do not contradict the conclusion that this interaction simply represents the expected tendency for deletion to become less important as redundancy is increased. Although conclusions about the primary variable are restricted by the presence of an interaction, it appears that reaction time is longer for more redundant stimuli.

Stimulus redundancy and incomplete stimuli.—Each of the 45 *Ss* in Cond. 10%-I was presented with 20 incomplete stimuli distributed throughout his 200 trials. Table 6 shows the per cent correct responses, for each condition of stimulus redundancy within Cond. 10%-I, to incomplete stimuli. The number correct was corrected for chance by the appropriate formula for each condition of response uncertainty, as previously explained. Both the four-element and five-element groups made a significantly higher percentage of correct responses to incomplete stimuli than did the three-element group. This result confirms the prediction that redundant stimuli would be easier to identify than nonredundant stimuli when there was one element missing.

TABLE 6

PER CENT CORRECT RESPONSES OF THE 10%-I GROUP TO INCOMPLETE STIMULI AVERAGED FOR ALL CONDITIONS OF RESPONSE UNCERTAINTY

Measure	Number of Elements		
	3	4	5
P	20.8	52.8	36.8
SE_P	2.34	2.88	2.72

Note.—The percentages are based on the 20 incomplete stimuli presented to each of the 45 *Ss* in the 10%-I group.

There are two apparent irregularities in Table 6. First, it may seem surprising that accuracy on incomplete stimuli was not zero for the three-element group, since the scores are corrected for chance and with one element deleted a three-element stimulus cannot be perfectly identified. There are two elements left in incomplete three-element stimuli, however, and these cues make it possible to narrow down the number of stimuli which could have been presented to two. In other words, *Ss* do better than chance on incomplete three-element stimuli because some information is still supplied by the stimulus. The other irregularity is the fact that the four-element group is significantly more accurate than the five-element group on incomplete stimuli. The data of the present experiment provide no convincing explanation for this result, so it must go without elaboration in this account. It does not alter the fact, however, that redundant incomplete stimuli were easier to identify than nonredundant incomplete stimuli.

Response uncertainty and the course of learning.—The per cent correct for 2, 4, and 8 response categories was 71.7, 78.6, and 70.7, respectively. The accuracy scores were corrected for chance in the usual manner. The differences are not highly significant, and there is no evidence for a consistent effect of response uncertainty on the rate of learning in this situation. This may be interpreted to mean that the perceptual task does not seem to be affected by changes in the response requirements.

Response uncertainty and reaction time.—Mean reaction time (in sec./100) over the last 50 trials, for the 2, 4, and 8 response categories, was

77.6, 85.0, and 121.4, respectively. The analysis of variance shows response uncertainty to be a significant source, free from interactions. The data support the conclusion that reaction time is an increasing function of response uncertainty.

Discussion

Formal alternative-cue redundancy clearly did not facilitate the identification of stimulus patterns in this experiment. This study, however, does not answer the question of whether redundancy *can* be helpful. Rather, it shows that one possible result of making stimuli logically redundant is to make them more difficult to identify. In other words, stimulus redundancy seemed to increase *S*'s uncertainty about the identity of the stimuli rather than to decrease it.

To understand this effect, it is necessary to examine the operation which was performed on the stimuli in introducing what is called stimulus redundancy. This operation was the addition of one or two binary elements to the adequate three-element stimuli. To be sure, this operation had the effect of making the stimuli redundant in a formal sense, but it also changed the form of the patterns defined by the lighted glow tubes. Since the stimuli were apparently perceived as whole patterns, this latter effect has the greater significance for the relative difficulty of identification under the various conditions of stimulus redundancy. It was shown that extra elements made the stimuli more difficult to identify a priori, in terms of interpattern relationships, and that the data bore out the predictions based on this kind of stimulus specification. These findings indicate that the four- and five-element stimuli were redundant in a formal sense only, and were not really perceptually redundant. It is conceivable that stimuli which are perceptually redundant, i.e., optimally coded, might produce quite different results.

Reaction time was an increasing function of response uncertainty alone when the stimulus conditions were held constant. The data suggest that the perceptual task performed by *S* also remained constant as response uncertainty was varied, even though the amount of information logically necessary for stimulus identification increases with response uncertainty. This conclusion is suggested by the fact that learning proceeds at about the same rate under all conditions of response uncertainty. In other words, the relationship between reaction time and response uncertainty found in this experiment appears to be independent of stimulus uncertainty.

Hyman (2) has described a relationship between response uncertainty and reaction time in an experiment in which stimulus uncertainty and response uncertainty were always equivalent. His data and those of this experiment are not strictly comparable, for several reasons. Hyman obtained his results under conditions of perfect accuracy, long practice, and a quite accurate estimate of the uncertainty of each condition. The functions from the two experiments may still be compared, however, if it is assumed that 2, 4, and 8 response categories represent approximately 1, 2, and 3 bits of response uncertainty,

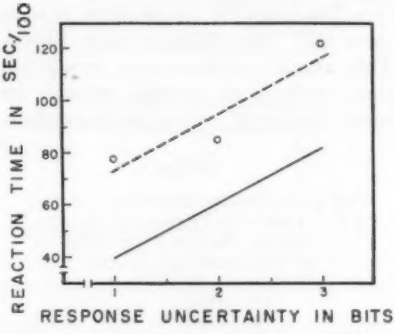

Fig. 4. Reaction time as a function of response uncertainty. The open circles are the data from this study, and the broken line is the best fitting straight line through those points. The solid line is the regression of reaction time on stimulus information for one *S* in an experiment by Hyman (2).

respectively. This assumption permits the plotting of the reaction times obtained in the present study as a function of response uncertainty in bits, as in Fig. 4. The best fitting straight line has been drawn through the three points from this experiment, and Fig. 4 also includes the regression line for one of Hyman's *S*s. The slope of the function for Hyman's *S* is no greater than that of the function from the present study, so it appears that reducing stimulus and response uncertainty together decreases reaction time no more than reducing response uncertainty alone. Comparison with the function for any of Hyman's *S*s does not materially change this conclusion. There is a possibility that further practice could reduce reaction time for the higher uncertainty conditions, and thus decrease the slope of the functions from this experiment. But as it stands, this comparison suggests that changes in response uncertainty alone account for most of the slope of the function relating reaction time to the covariation of stimulus *and* response uncertainty. Stimulus factors, within the range investigated, appear to affect only the intercept constant of the function.

Summary

The purpose of this study was to observe the effect of two major variables, stimulus redundancy and response uncertainty, on the process of identifying stimuli composed of binary elements. The rate at which a list of eight such stimuli was learned and the reaction time required for identifying the stimuli were used as measures of the identification process. Stimulus redundancy was varied by adding elements to the minimum number of elements necessary to identify each stimulus. This theoretically provided *S* with alternative ways to identify the stimuli. Response uncertainty was varied by changing the number of responses available to *S*, thereby changing the number of stimuli to which each response referred. A third variable consisted of deleting one element from 10% of the stimuli presented to half of the *S*s. This deletion variable was introduced to test the effect of stimulus redundancy on the accuracy of responses to incomplete stimuli.

Stimulus redundancy retarded learning, by making the stimulus patterns more complex and difficult to identify, and slightly increased reaction time. In this form, redundancy cannot be considered an important aid to stimulus identification. The *S*s using redundant stimuli, however, were more accurate on incomplete stimuli than those whose stimuli were nonredundant.

Response uncertainty had no consistent effect on the rate of learning, but was related positively to reaction time. The relative importance of stimulus and response factors in this relationship is discussed.

References

1. Gibson, E. J. Intra-list generalization as a factor in verbal learning. *J. exp. Psychol.*, 1942, **30**, 185–200.
2. Hyman, R. Stimulus information as a determinant of reaction time. *J. exp. Psychol.*, 1953, **45**, 188–196.
3. Saltzman, I. J., & Garner, W. R. Reaction time as a measure of span of attention. *J. Psychol.*, 1948, **25**, 227–241.

(Received May 24, 1954)

Journal of Experimental Psychology
Vol. 49, No. 2, 1955

RELATION OF STIMULUS AND RESPONSE AMPLITUDE TO TRACKING PERFORMANCE [1]

BRYCE O. HARTMAN [2] AND PAUL M. FITTS

The Ohio State University

The relation of stimulus magnitude to accuracy of estimation and to the precision of motor response has long been a topic of interest in psychophysics. Relatively little attention has been given, however, to the relation of stimulus and response amplitudes to performance in continuous tasks such as tracking. In a typical tracking task *S* operates a control and observes a separate visual display. Such an arrangement permits independent experimental manipulation of stimulus and response amplitude. These variables will be referred to hereafter as display (S_d) and control (S_c) scale factors and their effects on tracking performance as scale effects.

The most comprehensive studies conducted to date on scale effects in continuous tasks are Helson's (**5**) wartime investigations of handwheel tracking. One of Helson's most general conclusions, stated as the **U** hypothesis, was that relative performance tended to be maximum over a wide range of task conditions, such as speed and amplitude of turning movements or mass of the handwheel control. This type of relation is similar to the familiar Weber-type function found in many psychophysical studies using discrete stimuli and responses. The finding is also in general agreement with results reported by one of the present writers (**3**) on the interrelations among the amplitude, accuracy, and rate of responding in serial motor tasks. In spite of the basic agreement among these varied types of data, however, there are grounds for expecting that the precise relation of stimulus and response amplitude to performance in continuous tasks may depart considerably from a Weber-type function in certain amplitude regions, and that the relation may vary considerably with different secondary task conditions. Important task conditions, for example, are the degree of patterning or predictability of the motor response sequence, and the relative importance of visual vs. proprioceptive sources of information.

The present investigation is best described as consisting of a series of self-contained experiments, in each of which the level of performance was determined for 25 combinations of display and control scale factors for a given set of secondary task conditions. Several types of target motion, differing in frequency and intricacy, and both compensatory and pursuit displays were used. Highly practiced *Ss* were employed, but one task condition was repeated at the beginning and end of the 11-session training period and again at the end of the series of experiments as an empirical check on practice effects.

[1] This research was supported in part by the United States Air Force under Contract No. AF 33(038)-10528 with The Ohio State University Research Foundation, monitored by the Air Force Personnel and Training Research Center. Permission is granted for reproduction, publication, use, and disposal, in whole or in part by or for the United States Government. Some of the data here reported are included in a Ph.D. dissertation submitted by the first author to the Graduate School of The Ohio State University.

[2] Now Captain, MSC, USA, Army Medical Research Laboratory, Ft. Knox, Kentucky.

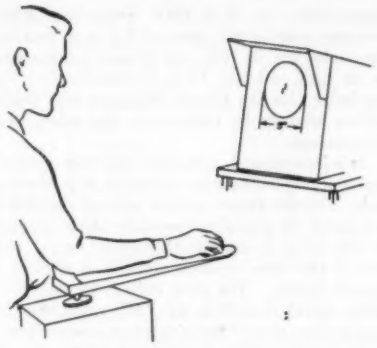

FIG. 1. Sketch of the apparatus showing the spatial arrangement of the display and the arm control.

METHOD

Apparatus.—A two-channel electronic tracking apparatus, whose detailed construction and functional characteristics have been described elsewhere (14), was used. Its principal components were (*a*) a target generator, (*b*) a programming unit, (*c*) duplicate cathode ray tube (CRT) displays, (*d*) duplicate tracking controls, and (*e*) duplicate scoring and recording units.

The apparatus permitted two *S*s working in separate test booths to be tested simultaneously. Both *S*s tracked the same problem, but each saw only the results of his own responses. Each tracking booth contained a voice communication system for giving instructions.

The CRT display was 5 in. in diameter, 28 in. away and 15° downward from *S*'s eyes, and perpendicular to the line of sight. The target was a bright vertical line $\frac{1}{16}$ in. wide and $\frac{1}{4}$ in. long. The cursor was a vertical line of similar dimensions which appeared below and slightly overlapped the target line. The target and cursor lines were painted on the display by a single electron beam which was deflected alternately by the two signals operating through a 60-cps time-sharing switch; this procedure eliminates problems of drift and parallax between display and scoring circuits. Target and cursor moved along a horizontal path which bisected the CRT. The *S*'s task was to keep target and cursor in continuous vernier alignment. The level of illumination in the booths was low (.1 ft.-c.) in order to provide good contrast on the CRTs.

The control (see Fig. 1) was a horizontal lever.[3] The *S* placed the elbow of his right arm above its pivot point and extended his arm forward, grasping a knob at the end of the lever. The position of the knob was adjusted to match the length of *S*'s arm.

Secondary task variables.—Two types of displays were employed, pursuit and compensatory. Functional block diagrams of the apparatus arranged for each type of display are given in Fig. 2. With a pursuit display the motion of the target is determined by variations in the input voltage (θ_i) of the target course generator, and the motion of the cursor is determined by variations in the output voltage (θ_o) from *S*'s tracking control. The instantaneous magnitude of tracking error is the difference between input and output voltages. For any given values of S_d and S_c this error voltage bears a fixed relation to the horizontal distance

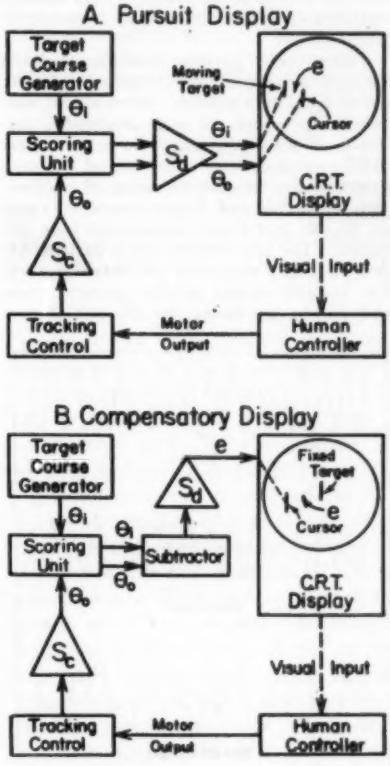

FIG. 2. Block diagram of the electronic tracking apparatus arranged for a pursuit display (A) and a compensatory display (B).

[3] Data were obtained with two other types of controls, a joy stick and a wrist control, but this task variation produced no significant variation in scale effects, and the data are not included in the present report.

in inches between target and cursor, and to the angular difference in degrees or in mils between the existing and the correct position of the arm control.

When the compensatory condition is used, the target line remains stationary in the center of the display and only the cursor moves. The voltages from the target course generator and from *S*'s control are fed to a subtractor which is inserted ahead of the display, and the difference between these two voltages, i.e., error, appears on the CRT display as cursor position.

With both types of displays the same difference in voltages is used to activate the scoring circuits. With both displays the tracking control must be moved to the right to introduce a correction to the right. Exactly the same motor response in relation to a given magnitude and direction of error is required for perfect tracking regardless of whether the display is pursuit or compensatory.

Target motion patterns consisting of simple harmonic motion of three frequencies, 15, 30, and 60 cpm, were studied. These signals were identical in respect to peak amplitudes, amplitude distributions, and root mean square (RMS) voltages over a sample period of 1 min. Target motion patterns consisting of combinations of (*a*) 30- and 20-cpm components and (*b*) 30-, 20-, and 10-cpm components were also studied. The two complex target patterns are shown in Fig. 3 along with the 30-cpm pattern. The complex target motion patterns were produced by combining the 30- and 20-cpm components so that their respective *RMS* voltages were in the ratio of 3.1 to 1, and by combining the 30-, 20-, and 10-cpm components in the ratio of 1.6 to 2.5 to 1, respectively. In the latter case the 10-cpm frequency was phase shifted slightly in relation to the other two components.

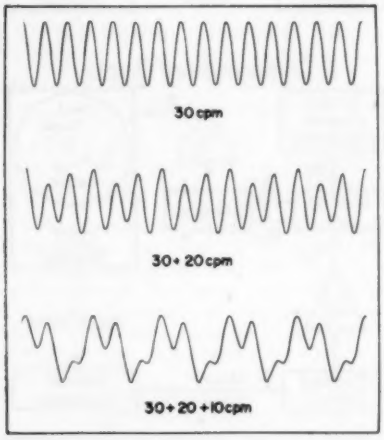

FIG. 3. Thirty-second graphic records of target motion patterns of three levels of intricacy: a 30-cpm sine wave, a combination of 30 + 20-cpm sine waves, and a combination of 30 + 20 + 10-cpm sine waves.

It is important to note that when two or more time-varying voltages are combined to produce a more intricate target motion pattern the *RMS* voltage of the resulting composite signal around its own mean is equal to the square root of the sum of the mean squared voltages of the component signals. The peak voltage of the composite signal as well as its *RMS* value thus is greater than that of any single component; however, the amplitude distribution of the signal changes in such a way that the increase in peak voltage is greater than is the increase in *RMS* voltage. In order to provide a basis of comparison between the three simple and the two more intricate target motion patterns studied in this experiment, the voltages of the composite signals were reduced until the peak plus or minus voltage, which corresponded to the maximum displacement of the problem cursor to left or right of its center position, was the same for all five target courses. The relative *RMS* voltages of the two complex signals, in order of increasing complexity, were 79% and 75%, respectively, of that of the 30-cpm signal.[4]

Display and control scale factors.—S_d is defined as the amplitude in inches traversed by the target on a pursuit display. With a compensatory display the error cursor traverses this same range if the control is held in its neutral position, i.e., if error corresponds to target course input. The five categories of S_d studied were ¼, ½, 1, 2, and 4 in. Linear distances of these amounts on the CRT display are equivalent to visual angles ranging from approximately 30 sec. to slightly over 8° at *S*'s eye point.

S_c is defined as the range of angular motion of the control required to perform the tracking task perfectly. The five categories of S_c studied were 5°, 10°, 20°, 40°, and 80° of arm rotation at the elbow. Stops prevented the arm control from being rotated over a range of more than 90°. For convenience the 5° condition will be referred to as the smallest control scale factor and the 80° condition as the largest one. The control is of course less sensitive, i.e., the ratio of display motion to control motion is smallest in the latter case.

Selector switches at *E*'s console permitted

[4] An alternative procedure would have been to equate the *RMS* values and permit the ranges of target motion to vary for the different courses.

TABLE 1

SEQUENCE OF TRACKING TASKS

Sessions	*Ss* 1–4	*Ss* 5–8
1–3	30 cpm, *P*	30 cpm, *P*
4–11	Training	Training
12–15	30 cpm, *P*	30 cpm, *P*
16–18	15 cpm, *P*	60 cpm, *P*
19–24	30 cpm, *C*	30 cpm, *P**
25–27	30+20 cpm, *C*	30+20 cpm, *P*
28–30	30+20+10 cpm, *C*	30+20+10 cpm, *P*
31–32	30 cpm, *P*	30 cpm, *P*

Note.—*P* represents a pursuit display, *C* a compensatory display. During Sessions 4-11 practice was given on each of the tracking tasks.

* During these sessions two new types of controls were used.

changes in the values assigned to S_d and S_c to be made between trials.

Procedure.—Eight volunteer college men served as *Ss*. All were right handed. They were paid by the hour. Sessions consisted of 25 trials. Each trial was 90 sec. long but only the last 60 sec. were scored. A rest period of 3 min. occurred after Trials 5, 10, and 20, and one of 10 min. occurred after Trial 15. During these rest periods *Ss* relaxed outside their booths. A 1-min. rest period was given between all other trials, *Ss* remaining in their booths. A buzzer sounded before the beginning of each new trial. Another buzzer, which remained on during the trial, effectively masked room noises and possible auditory cues from the apparatus. All timing and scoring was automatic. The *Ss* alternated booths on successive sessions. Eleven 1.5-hr. sessions were devoted to training. During the first three of these sessions a standard (30-cpm pursuit) task was used; this task constituted the first task of the series of experimental sessions and was repeated on Sessions 31 and 32 as an empirical check on learning effects. During Sessions 4 through 11 pretraining was given on all of the task conditions employed in later experimental sessions.

Each pair of *Ss* was first scheduled for a demonstration session. On each of the subsequent 32 training and test sessions each *S* was tested on each of the 25 conditions of S_c and S_d. Three different balanced sequences of the scale factors were used. The *Ss* were required to adjust to the scale factor changes during the initial 30 sec. of each trial, which were not scored.

During the block of sessions assigned to each experiment the type of tracking task was held constant. The sequence of secondary task conditions is given in Table 1. The *Ss* were fully informed of these constant conditions at the start of each session.

Performance criterion.—The measure of performance on each trial was a time-on-target score, the time during the 60-sec. scoring period that the magnitude of the error was equal to or less than an arbitrary reference voltage. The critical error voltage for obtaining a time-on-target score was .325 v. or 5% of the peak plus or minus input voltage from the problem generator (which varied between plus and minus 6.5 v.). This criterion bears a fixed relation to any particular value of S_c or S_d, and the equivalent absolute error with respect to the stimulus or the response can be determined easily. For the smallest display scale of $S_d = \frac{1}{4}$ in. (or $\frac{1}{8}$ in. to either side of the center of the display), an error of 5% is equivalent to .00625 in. or .75 min. of visual angle to either side of the target; for the smallest control scale of $S_c = 5°$, an accuracy of $\pm$ 7.5 min. or $\pm$ 4.5 mils in positioning the control is required in order to remain on target. As long as tracking accuracy varies in direct proportion to the display or control scale factor, as would be the case if the relation of error to stimulus and response amplitude followed a Weber-type function, time-on-target scores should remain constant as scale factors are changed.

RESULTS

Tracking performance data for eight different secondary task conditions and for one condition at two stages of practice are shown graphically in Fig. 4. The parameter for each family of curves is the value of S_c. Time-on-target scores are given on the *y* axis and steps of S_d on the *x* axis. Each point is based on the pooled data for all *Ss* who worked at that set of conditions, and for the number of sessions indicated in Table 1.

Visual and motor scale effects.—In Fig. 4 the degree of separation of the curves within each family reflects the departure of the control (S_c) scale effect from a Weber-type relation between error and response amplitude, and the slope of each of the curves indicates the departure of the display (S_d) scale effect from a proportional relation between error and visual stimulus amplitude. In other words, if the results of each experiment were consistent with a Weber-type function, the data for each graph in Fig. 4 would fit a single horizontal straight line. Obviously such is not the case.

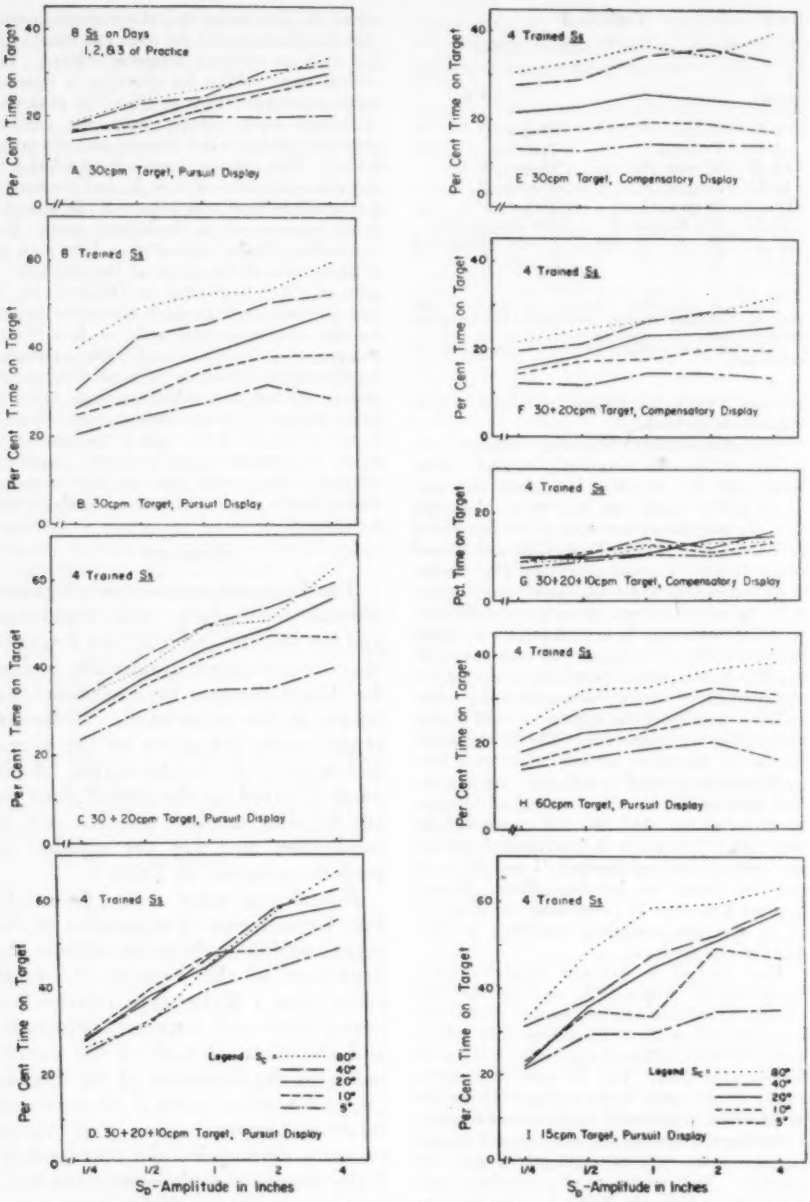
8 Ss on Days 1, 2, &3 of Practice
A 30cpm Target, Pursuit Display
4 Trained Ss
E 30cpm Target, Compensatory Display
8 Trained Ss
B 30cpm Target, Pursuit Display
4 Trained Ss
F 30+20cpm Target, Compensatory Display
4 Trained Ss
G 30+20+10cpm Target, Compensatory Display
4 Trained Ss
C 30 + 20cpm Target, Pursuit Display
4 Trained Ss
H 60cpm Target, Pursuit Display
4 Trained Ss
D 30+20+10cpm Target, Pursuit Display
Legend S_c = 80° 40° 20° 10° 5°
4 Trained Ss
I 15cpm Target, Pursuit Display
Per Cent Time on Target
Pct. Time on Target
S_d-Amplitude in Inches
1/4 1/2 1 2 4

FIG. 4. Time-on-target scores as a function of control (S_c) and display (S_d) scale factors for nine experimental tracking tasks. The five categories of S_d are shown on the abscissa of each graph and the five categories of S_c are shown as separate curves. The number of Ss and number of sessions for each experimental condition are indicated in Table 1, and the significance levels of the S_d and S_c effects are indicated in Table 2.

Instead, time-on-target scores tend to improve as S_e or S_d values are increased. Scale effects can therefore be called *progressive*.

It should be recalled that the experimental design employed in this series of experiments was chosen to maximize the precision with which scale effects could be evaluated within each of the sets of secondary task conditions. Hence each set of data for a single set of secondary task conditions was subjected to a separate analysis of variance. The estimates of error used in computing the *F* ratios of (*a*) S_e effects, (*b*) S_d effects, and (*c*) $S_e \times S_d$ interaction were the (*a*) $S_e \times S$ interaction, (*b*) $S_d \times S$ interaction, and (*c*) $S_e \times S_d \times S$ interaction, respectively. The *F* ratios are given in Table 2. Data for the standard 30-cpm pursuit task are available for all eight *S*s at three stages of practice; in all other instances four *S*s were used.

An examination of Table 2 reveals that the display scale effect is significantly progressive at the .01 level of confidence for all nine sets of data, and the control scale effect is significantly progressive at the .01 level for all but one set of data.

The magnitude of the scale effects varied considerably depending upon the secondary conditions of the tracking task. Most of the remainder of the present section will be devoted to the examination of these differences. The basis for estimating the magnitude of scale effects is the value of ϵ^2 (**10,** Formula 220) for each set of secondary conditions. These values are included in Table 2. Although the magnitude of ϵ^2 is influenced by the arbitrary choice of categories for S_e and S_d (**8,** p. 63), the use of ϵ^2 as an indicant of the relative strength of the functional relations represented by different families of curves is considered justified by the present purpose since the sets of categories for S_e and for S_d cover a similar range (16 to 1) on a physical scale of magnitude, and since the comparisons of interest are always those between these particular categories as secondary task conditions are varied.

Interaction effect.—From an inspection of the data summarized in Fig. 4, it appears that relative tracking performance in most cases improves in an approximately linear manner with geometric increments of S_d, and that the improvement in time-on-target score with geometric increments in S_e is also approximately

TABLE 2

SUMMARY OF *F* RATIOS FROM SEPARATE ANALYSES OF VARIANCE ON EACH OF NINE EXPERIMENTAL TRACKING TASKS, AND CORRESPONDING VALUES OF ϵ^2

Tracking Tasks	*Ss*	*F* Ratio			ϵ^2		
		S_e	S_d	$S_e \times S_d$	S_e	S_d	$S_e \times S_d$
30 cpm, Pursuit (Days 1–3)	8	19**	59**	1.6	.69	.88	.07
30 cpm, Pursuit (Days 12–15)	8	144**	71**	3.3**	.95	.90	.22
30 + 20 cpm, Pursuit	4	86**	59**	1.7	.96	.94	.14
30 + 20 + 10 cpm, Pursuit	4	6.5**	86**	4.6**	.58	.95	.47
30 cpm, Compensatory	4	105**	7**	2.1*	.96	.60	.21
30 + 20 cpm, Compensatory	4	43**	23**	1.6	.91	.85	.13
30 + 20 + 10 cpm, Compensatory	4	1.7	24**	2.0*	.15	.85	.19
60 cpm, Pursuit	4	24**	14**	2.8**	.85	.76	.31
15 cpm, Pursuit	4	55**	100**	5.3**	.93	.96	.52

* Significant at the .05 level of confidence.
** Significant at the .01 level of confidence.

linear in the majority of cases. The basic relation between time on target (t) and the magnitude of either scale factor (S) over the range studied appears to be of the general form

$$t = A + B \log S.$$

Four of the nine $S_c \times S_d$ interactions evaluated by the analysis of variance were significant at the .01 level of confidence. All four instances were conditions in which a pursuit display was used. It is apparent from Table 2 that the estimated magnitudes of the interaction effects for the categories of variables used in the present experiment in all cases except one are small in relation to the estimated magnitudes of the corresponding main effects. These results are in agreement with the findings from the two preliminary studies in this series (**4, 9**).

The interaction between S_c and S_d effects appears to arise at least in part from an increase in the progressive nature of the control scale effect as display magnification increases from $\frac{1}{4}$ to 4 in. Thus, it appears that as the visual display is magnified *Ss* found it increasingly advantageous to use a large response motion. This was especially true with a pursuit display.

Practice effects.—Data for an evaluation of practice effects are available for the 30-cpm pursuit task. This task was administered to all eight *Ss* during Sessions 1–3, 12–15, and 31–32. The extent of improvement from Days 1–3 to Days 12–15 can be seen by comparing Fig. 4A and 4B. As indicated by an F test, this improvement is significant at the .01 level of confidence.

A further analysis of the combined data for Sessions 1–3 and 12–13 was carried out in order to test for the presence of any changes in the magnitude of scale effects with practice. The test was made by means of an analysis of variance with sessions treated as an additional variable. Any significant change in the magnitude of either the visual or the motor scale effect as a function of the level of practice should appear as a significant interaction between training sessions and scale effects. The interaction of sessions with S_c (tested against the second-order interaction, Sessions $\times$ S_c $\times$ *Ss*) was found to be significant at the .01 level of confidence. The interaction of sessions with S_d (tested by the second-order interaction, Sessions $\times$ S_d $\times$ *Ss*) was found to be significant at the .05 level. Thus, the quantitative results support the conclusion that in this task the magnitude of scale effects increased from the beginning to the end of the series. As evaluated by ϵ^2 this interaction with sessions was considerably larger in magnitude for S_c than for S_d.

No significant change in performance occurred between Sessions 12–15 and Sessions 31–32. This result lends support to the assumption that the 275 90-sec. practice trials (Sessions 1–11) were sufficient to bring *Ss* to a relatively stable performance level. Therefore, although the experimental design was chosen to facilitate analyses of scale effects within a series of self-contained experiments, comparisons of the magnitude of scale effects among different secondary conditions appear to be justified. For the purpose of such a comparison it is important to note that no significant differences existed between *Ss* 1–4 and *Ss* 5–8 either at the beginning or at the end of the series of experiments.

Target frequency.—The effect of target frequency, for simple harmonic target motion, is revealed in Fig. 4I, 4B, and 4H. Performance was best at 15 cpm, almost as good at 30 cpm, and poorest at 60 cpm. At the highest frequency (the most difficult condition) the advantage of visual magnification was reduced; in par-

ticular, no improvement resulted from increasing S_d beyond 2 in. of magnification.

Intricacy of target motion.—Performance with a compensatory display became progressively poorer as target motion was made progressively more intricate. However, an opposite change in performance resulted from an increase in target course complexity for *S*s using a pursuit display. Performance actually improved slightly, although not significantly so, with each increase in intricacy of target course. In interpreting these results it should be recalled that intricacy was produced by adding lower frequency components to a basic 30-cpm signal. The addition of lower harmonics was accompanied by a reduction in over-all signal power and also resulted in a shift of the mean power of the input signal toward lower frequencies, i.e., a reduction in the average rate of target motion. Thus, although a complete cycle of target motion consisted of a more intricate sequence, target motion was slower. These two factors, signal power and intricacy of wave form, apparently affect tracking performance in a markedly different way, depending on whether the pattern of target motion or only tracking error can be observed directly on the tracking display.

The influence of the intricacy of target motion on control scale effects differed markedly from its influence on display scale effects. On the one hand, an increase in intricacy resulted in a marked reduction in the S_c effect as indicated by a drop in the magnitude of ϵ^2. On the other hand, the S_d effect tended to increase slightly with target intricacy for both pursuit and compensatory conditions.

The significance of the change in the magnitude of visual and motor scale effects as a function of the intricacy of target motion can be evaluated by testing the significance of the interaction term in the analysis of variance across levels of complexity. Two such analyses were carried out, one for the four *S*s who were tested under three degrees of complexity with a pursuit display, and one for the four *S*s who were tested under the same three degrees of complexity with a compensatory display. The error term used to evaluate the significance of the interaction was the second-order interaction, *S*s × complexity × the scale effect in question. For the pursuit display the interaction of target course complexity with both S_d and S_c was significant at the .01 level of confidence. For the compensatory display the interaction of target intricacy with S_c was significant at the .01 level of confidence, but the interaction of intricacy with S_d was not significant.

Discussion

The present investigation covers a wide range of both scale factor variables and secondary task variables. Nevertheless, the reader should keep in mind that all types of target motion studied were cyclical and therefore predictable, that no frequencies outside the band from 10 to 60 cpm were employed, and that only a displacement tracking task was studied. In spite of such limitation in the range of secondary variables, however, the findings with regard to visual and motor scale effects appear to support a number of rather general inferences as to the nature of tracking behavior.

The finding of greatest generality is that an increase in visual stimulus or in motor response amplitude is accompanied by a reduction in the ratio of error to stimulus and response amplitude, respectively. Thus the results do not follow a Weber-type function. Several probable bases for these progressive scale effects can be suggested.

The visual scale effect.—In order to track a moving target with small error *S* must make accurate discriminations of stimulus magnitude and of stimulus rate. Relative accuracy in judging error magnitude should improve as the mag-

nification of the distance between target and cursor is increased from the threshold up to some limit. Also, as Senders (12) has pointed out, the use of a frame of reference improves tracking performance. In the present tracking task the 5-in. diameter rim of the CRT display served as such a reference. The value of this frame of reference should increase as the limit of cursor motion approaches the rim of the CRT display, i.e., as S_d is increased.[5]

Effective tracking performance with a simple displacement control system and a coherent target course requires the discrimination of rate as well as position information (13), since rate information provides a basis for predicting future target position. Cursor rate increases as S_d is increased, and rate discrimination should improve with S_d, up to some limit. There is no reason to believe, however, that the value of S_d that is optimum for the discrimination of rate should necessarily be optimum for the discrimination of position. An optimum degree of display magnification probably is one that affords a good compromise for the joint discrimination of position and rate information, given a particular type of target motion.

Pursuit vs. compensatory displays.—As S improves his performance with a pursuit display he continues to have available complete information regarding target rates and movement patterns. However, the better his performance with a compensatory display the less information he receives, since with perfect tracking the cursor as well as the target remains motionless in the center of the display. Such differences have been treated in detail by other writers (1, 12). The only apparent advantage of a compensatory display is the greater magnification of error that is possible. The present results indicate, however, that tracking performance with a pursuit display is very much better, especially for complex but periodic target courses, than is performance with a compensatory display, even when a 16 times greater stimulus magnification is used with the latter display. When S uses a compensatory display the only way that he could learn to predict target motion is by continually comparing the proprioceptive feedback arising from his own movement responses with visually observed error information, and by inferring target motion to be the difference between these two sensory inputs. This is probably an extremely difficult analysis. Since the performance even of trained Ss is so greatly facilitated by the use of a pursuit display, it appears possible that for predictable target courses initial training with a pursuit display might result in much greater positive transfer to a compensatory display (the kind of display used in many military and industrial tasks) than would initial practice on the much more difficult compensatory display itself.

The motor scale effect.—Relative improvement in the accuracy of motor responses with an increase in the range of the forces applied to a control has been noted by previous investigators. Jenkins (7) reported that the ratio of variable error to the magnitude of the force applied to an isometric control in a discrete task decreased up to at least 10 lb. of force, while Helson (5) found that the relative error in tracking both uniform- and variable-rate target courses decreased as the maximum required speed of handwheel turning increased to a point near the maximum rate at which S was capable of rotating the control (about 200 rpm). These results are in agreement with the present findings for the S_e effect. It appears, therefore, that the motor system with its associated proprioceptive feedback loops is utilized most effectively when the full range of S's output capabilities is demanded by the task rather than when only a restricted range of forces or movement amplitudes is required. The interpretation of these effects is not a simple matter, however. An increase in S_e results in complex changes in both the physical and the informational aspects of the control process.

The S_e effect was most marked for the simplest target courses and least marked for the most intricate signal. There is no reason to believe that the motor

[5] An alternative procedure would have been to reduce the size of the CRT display in proportion to the reduction in S_d from 4 in. to $\frac{1}{4}$ in.

system should experience greater difficulty, on physical grounds alone, in producing the more intricate motion pattern since the average rate of motion, and therefore the energy required to produce the motion, is lower than for a single frequency course of the same amplitude. It is tentatively concluded, therefore, that the reduction in the advantage normally associated with a large value of S_e resulted from the fact that *S*s found it difficult to learn to utilize proprioceptive cues when executing the more intricate movement patterns. If this reasoning is correct the S_e effect might be expected to continue to become increasingly progressive for long periods of practice on a complex but predictable target course.

Jenkins and Olson (6) have reported an optimum ratio of control to display movement, in the case of discrete control adjustments, and Helson's (5) data also indicate an optimal range. No such clear maxima were found over the range of primary and secondary task conditions used in the present studies of continuous movements, nor do the results offer much support for the U hypothesis. The only exceptions are an apparent leveling off of the progressive nature of the S_e effect at about 40° of arm movement for the more intricate target courses tracked with a pursuit display, and a possible leveling off of the S_d effect at 2 in. of visual display diameter for the highest (60-cpm) frequency.

Shift from visual to proprioceptive control of movement sequences.—Judged in terms of the magnitude of visual and motor scale effects the present findings indicate a shift in the relative importance of visual as contrasted with proprioceptive or motor factors as the level of learning and as secondary task conditions changed. Optimum visual stimulus magnitude appears to be more important than optimum response magnitude (or optimum proprioceptive feedback) for complex target courses and for early stages of practice. In contrast, optimum response magnitude appears to be more important than optimum visual conditions for simple, highly practiced movement patterns. It is hypothesized that this effect reflects a change in the relative importance of visual versus proprioceptive information for the control of relatively simple but highly learned sequences of movements.

Various writers have postulated a gradual shift from visual to proprioceptive control as *S*s gain increasing proficiency in a motor task, but the dependence of such a shift on the intricacy of the response sequence has not been emphasized. However, it appears reasonable that proprioceptive feedback arising during a movement response should be utilized most effectively when the movement pattern is simple and highly repetitive. Some evidence for this hypothesis is furnished by the finding that simple and highly practiced sequential movements, such as those made in walking, rowing, or tapping at a preferred rate, exhibit extremely small variability in temporal and spatial pattern (2, 11, 15).

In apparent disagreement with the generalizations drawn from the present study is Helson's finding (5) that the beneficial effect of large handwheel mass and high rates of handwheel turning were largest early in practice. However, the conditions studied by Helson included response rates almost seven times higher than those used in the present study, a more complex target motion, and the use of a circular rather than a reciprocal arm motion. The net effect of these secondary conditions may have been to make Helson's task one in which *S*s were unable to learn repetitive movement patterns, and thus one in which a shift from visual to proprioceptive cues was not possible.

Summary

The present study investigated the effects of visual stimulus and motor response amplitude and their interactions in a continuous tracking task. Determinations of scale effects were made for eight secondary task conditions including both pursuit and compensatory tracking displays, three categories of intricacy of the target course, and three frequencies of target motion.

With only one exception relative performance improved significantly as display amplification was increased from ¼ to 4 in. and as the amplitude of required arm control motion was increased from 5° to 80°. Thus, both scale effects were progressive and a Weber-type function did not

hold. However, the extent of this effect varied markedly with different tracking tasks.

As performance improved with practice on the standard task (30-cpm target motion with a pursuit display) the use of larger amplitudes of movement and larger display magnifications became increasingly beneficial. Larger amplitudes of control motion also became increasingly beneficial as the visual display scale was increased.

As the target motion was made more intricate by the addition of lower harmonics (and concomitantly was reduced in average power) the advantage of large amplitude control movements was reduced; however, the beneficial effects of display magnification increased slightly with increasing intricacy of target motion. The over-all performance level was severely reduced when a compensatory display was used and the target course was made more intricate, whereas when a pursuit display was used the more intricate but somewhat slower target motion was no more difficult to track. Pursuit tracking was markedly superior to compensatory tracking and this superiority was greatest for the most intricate pattern of target motion.

As the frequency of the single-component target motion was increased from 15 to 30 to 60 cpm, the progressive nature of the visual scale effect decreased. Rates of target motion in the region of 60 cpm appear to exceed an optimum discriminatory range when large S_d values are used.

Although visual and motor scale effects could be varied independently, the present experimental procedure did not permit the independent study of the effect of variations in proprioceptive feedback apart from the effect of a changed activity (energy) level in the motor system. The relations among the different sets of data, however, provide a basis for inferences regarding the relative value of position vs. rate information, and the relative value of visual vs. proprioceptive cues, in a tracking task. Changes in visual and motor scale effects are postulated to accompany a shift from primary dependence on visual cues to increased dependence on proprioceptive information in controlling sequences of movements.

References

1. Birmingham, H. P., & Taylor, F. V. A human engineering approach to the design of man-operated continuous control systems. Washington, D. C.: Naval Research Laboratory, 1954. (*NRL Rep.* No. 4333.)
2. Feige, K. Präzisionsleistung menschlicher Motorik. *Z. angew. Psychol.*, 1934, **69**, 1–88.
3. Fitts, P. M. The information capacity of the human motor system in controlling the amplitude of movement. *J. exp. Psychol.*, 1954, **47**, 381–391.
4. Fitts, P. M., Marlowe, E., & Noble, M. E. The interrelations of task variables in continuous pursuit tasks: I. Visual display scale, arm control scale, and target frequency in pursuit tracking. *USAF, Hum. Resour. Res. Cent., Res. Bull.*, 1953, No. 53–34.
5. Helson, H. Design of equipment and optimal human operation. *Amer. J. Psychol.*, 1949, **62**, 473–497.
6. Jenkins, W. L., & Olson, M. W. The use of levers in making settings on a linear scale. *J. appl. Psychol.*, 1952, **36**, 269–271.
7. Jenkins, W. O. The discrimination and reproduction of motor adjustments with various types of aircraft controls. *Amer. J. Psychol.*, 1947, **60**, 397–406.
8. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* New York: Houghton Mifflin, 1953.
9. Noble, M. E., Fitts, P. M., & Marlowe, E. The interrelations of task variables in continuous pursuit tasks: II. Visual display scale, arm control scale and target frequency in compensatory tracking. *USAF, Hum. Resour. Res. Cent., Res. Bull.*, 1953, No. 53–55.
10. Peters, C. C., & Van Voorhis, W. R. *Statistical procedures and their mathematical bases.* New York: McGraw-Hill, 1940.
11. Rethlingshafer, D. Measurement of a motor set. *J. exp. Psychol.*, 1943, **32**, 75–81.
12. Senders, J. W. The influence of surround on tracking performance: I. Tracking on combined pursuit and compensatory one-dimensional tasks with and without a structured surround. *USAF, WADC Tech. Rep.*, 1953, No. 52–229.
13. Walston, C. E. A mathematical analysis of the human operator in a closed-loop control system. Unpublished doctor's dissertation, Ohio State Univ., 1953.
14. Warren, C. E., Fontaine, A. B., & Clark, J. R. A two-dimensional electronic pursuit apparatus. *USAF, Hum. Resour. Res. Cent., Res. Bull.*, 1952, No. 52–26.
15. Wirth, W. Die Zeitverhältnisse bei schnellem rechtshändigen Taktieren mit und ohne Nebenaufgabe der Rhythmisierung. *Arch. ges. Psychol.*, 1937, **98**, 245–278.

(Received for early publication November 22, 1954)

Journal of Experimental Psychology
Vol. 49, No. 2, 1955

COMPOUND TRIAL-AND-ERROR LEARNING AS A FUNCTION OF RESPONSE AVAILABILITY (N_R) [1]

CLYDE E. NOBLE

Louisiana State University

Among the variables which may determine performance in solving multiple-choice problems is the total number of responses available to the organism for its use. McGeoch (**3**), Melton (**4**), and Underwood (**11**) have referred to such a variable in terms of a "dimension of discovery," and Riley (**9**) recently demonstrated its influence in paired-associate verbal learning. The present investigation is concerned with extensive variations in degrees of freedom of action in a nonverbal trial-and-error learning situation employing human *S*s. The term *response availability* will be used to denote the number of possible all-or-none reactions in a given task or problem. Since behavior freedom is entirely under *E*'s control, availability is regarded as an independent rather than a dependent variable. For this concept we shall adopt the symbol N_R.

Trial-and-error learning refers to adaptive behavior in the classical, descriptive sense of the early comparative psychologists. Its objective characteristics have been formalized by Hull (**2**) and Spence (**10**). Problems of response selection in instrumental learning may appear in either the *simple* (single-link) or the *compound* (multiple-link) form. In the latter case, a series of basic S-multiple-R problems must be solved in a coordinated manner. Thus, the paradigm for compound trial-and-error learning is: $S_1 \rightarrow R_1$, $S_2 \rightarrow R_2$, . . . , $S_n \rightarrow R_n$. From rational considerations, the difficulty of learning an integrated behavior sequence should be some positive function of the number of alternative reactions (N_R) available at each choice point. Difficulty may also be contingent upon other variables, such as the distribution of practice, pacing or self-pacing, the number of links in the chain, the reinforcement schedule, the locus, delay, and duration of reinforcement, and the serial pattern of the required acts (homogeneity, alternation, simultaneity).

The present experiment will investigate paced four-link heterogeneous compound trial-and-error learning with immediate serial reinforcement under seven N_R conditions, holding task length constant. Of major interest will be (*a*) the difficulty-availability relationship, and (*b*) the nature of the basic acquisition curves for the various problem situations.

METHOD

Apparatus.—The learning device is called the Selective Mathometer. It is described in detail elsewhere by Noble and Farese (**8**). The apparatus consists essentially of a semicircular

[1] The experimental work for this study was performed as part of the United States Air Force Human Resources Research and Development Program. The opinions or conclusions contained in this report are those of the author. They are not to be construed as reflecting the views or indorsement of the Department of the Air Force.

This research was conducted in the Perceptual and Motor Skills Research Laboratory, Human Resources Research Center. S/Sgt. R. B. Card supervised the experiment. A portion of this paper was read before the Midwestern Psychological Association, April 1954.

array of 19 pushbutton reaction keys, among which *S* must select the correct responses in space and time. Any number of keys can be rendered unavailable by a set of removable individual wooden blocks which function as key covers. When a given key is uncovered it is considered an available response. Visual stimuli of any type may be presented in paced or self-paced sequences by electronically-controlled rotary slide projection. Exposure duration and dark period are independently variable, and a wide range of rates and task lengths is possible. Serial reward is given automatically by the onset of a green lamp (secondary reinforcement) on *S*'s response panel following each correct reaction. Delay and duration of reward are independently variable. Continuous graphic recording is provided for the time of occurrence, sequence, and duration of all responses in addition to the onset, duration, and rate of stimulus presentation. Magnetic counters also cumulate trial, correct, and error data.

The calibration constants in the present experiment were as follows: 2-sec. exposure duration, 1.44-sec. dark period, 0-sec. reward delay, and .75-sec. reward duration. The presentation rate was one slide every 3.44 sec. Erroneous responses, omissions, or late choices were simply not reinforced.

For this study it was desired that the stimuli have no intrinsic differential effects upon behavior. Accordingly, four paralogs were chosen from previous research in verbal learning which were known to be low in meaningfulness (**5**) and familiarity (**7**). Their average scale values were 1.41 and .67, respectively. These four items were prepared as 2 × 2-in. slides to constitute a single trial sequence. The inter-trial interval, composed of a blank and a cue slide, was 6.88 sec. Stimulus presentation followed an invariant, paced sequence for each experimental condition, but the correct responses were required in all four-factorial permutations.

Seven treatments were given, in which the ratio of the number of correct to available responses varied as follows: 4/4, 4/5, 4/6, 4/7, 4/8, 4/9, and 4/10. On *S*'s response panel the keys are numbered from left to right. Keys 4, 8, 12, and 16 were arbitrarily correct (R_+) for all seven N_R conditions, whereas the corresponding incorrect keys (R_-) were designated as follows:

N_R	Number of R_- Key
4	none
5	10
6	6 - - - - - - - - 14
7	6 - - - 10 - - - 14
8	2 - - - 6 - - - - - - - - 14 - - - 18
9	2 - - - 6 - - - 10 - - - 14 - - - 18
10	2 - - - 6 - - 9 - 11 - - 14 - - - 18

In each condition *S*'s problem on a given trial was to select from among the total possible choices the correct sequence of four responses to the four stimuli. The *S* was instructed to use the forefinger of his preferred hand in responding. Repeated choices were allowed for each stimulus, but *S* was informed that responses made during the dark periods would not be scored. The criterion of mastery was set at two successive correct trials without error.

Procedure.—The experimental plan utilized a simple randomized design, in which the interaction of treatment (availability) and sequential (permutation) effects was confounded with individual differences. Each of the seven treatments ($n = 48$) was administered under the 24 possible permutations of four reactions, with two *Ss* assigned without bias to each treatment-permutation combination.

The *Ss* consisted of 336 basic airmen at Lackland Air Force Base, sorted randomly into the various treatments within each permutation. They were tested individually, and none had prior experience with the apparatus. Following the experiment, *S* was isolated from the remaining pool of *Ss* in order to prevent artifacts due to discussion of the procedure.

Results and Discussion

The general results of the experiment are shown in Fig. 1. Successive correct first choices (R_+) are plotted as functions of mean trials for the seven conditions of response availability (N_R). The parameters denote the number of available responses. It is evident that N_R exerts a powerful

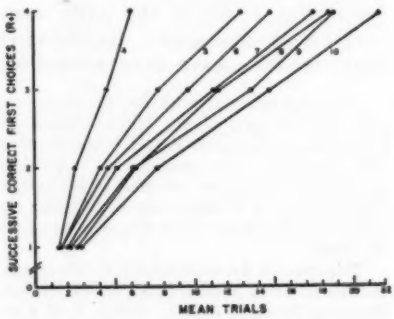

Fig. 1. Successive correct first choices as functions of mean trials for seven conditions of response availability.

retarding effect upon the rate of trial-and-error learning, in accordance with expectation.

From previous research by Noble (6) and Riley (9), it is known that difficulty and variability are positively correlated in tasks of this type. In view of the heterogeneity of variance among the treatment means of trials to mastery, Friedman's (1) nonparametric analysis of variance test was applied to the distributions of ranks among the experimental conditions. This analysis was carried out on a 7×24 matrix, in which each cell entry was a mean of two scores. For $df = 6$, $\chi^2 = 55.82$, which is highly significant ($p < .001$) of differences among the criterion performances. Difficulty is thus a positive function of N_R.

The graphical technique employed in Fig. 1 for summarizing the performance data is adequate for descriptive purposes, supplanting the laborious transformations of the Vincent method, but the measures do not possess any clear theoretical significance. Consequently, by analogy with simple conditioning, we turn to a consideration of probability of correct response (p_+) as a function of ordinal trials (N).

At the outset of compound trial-and-error learning, given a large number of counterbalanced observations, the theoretical probability of correct response per choice point ($1/N_R$) should be approximated empirically by the ratio of the frequency of correct (R_+) to total ($R_+ + R_-$) responses. A similar but more sensitive measure, and one somewhat easier to compute, is the relative frequency of correct intratrial *first* choices (p_+), considered as a proportion of the total possible first choices per trial (number of choice points $\times$ size of sample). This quantity may also be regarded as the average p_+ per choice point, particularly since the serial position error gradient in this investigation was approximately horizontal for all conditions. For present purposes, *Ss* reaching the criterion of mastery were assumed to continue functioning at $p_+ = 1.00$ to the limit of training required by the slowest *S*. Now by plotting p_+ as a function of N for selected values of N_R, it is possible to make a comparative study of the growth of trial-and-error learning proficiency in different tasks within a single rational coordinate system.

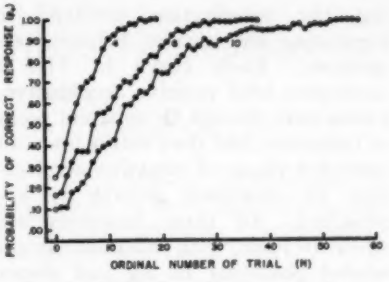

FIG. 2. Probability of correct first choice as functions of trials for three representative conditions of response availability (N_R). The circled origins are rational and are given by the reciprocal of N_R.

This analysis is presented in Fig. 2, where the probability of reaction evocation curves for the 4/4, 4/6, and 4/10 conditions has been isolated for analysis. The advantage of this measurement technique is at once evident: each performance curve originates at a hypothetical p_+ value which is indicative of the initial chance solubility of the problem. Thus, on the zero'th trial the p_+ values for 4, 6, and 10 available responses are the reciprocals .250, .167, and .100, respectively. These theoretical points align themselves appropriately with the subsequent empirical functions, incidentally verify-

ing the equalization involved in permuting the correct behavior sequences. Each curve in Fig. 2 undergoes brief positive acceleration, passes next through an apparent point of inflection, and then enters upon an extended phase of negative acceleration as maximum growth is approached. All three functions are ogival in form, with inflection points related positively to N_R and slopes negatively related.

Summary

The study involved paced four-link heterogeneous compound trial-and-error learning with immediate serial reinforcement under seven conditions of response availability (N_R) and task length held constant. Three hundred and thirty-six human *S*s were used in a simple randomized design and the ratios of correct to available responses were 4/4, 4/5, 4/6, 4/7, 4/8, 4/9, and 4/10. To a series of four stimuli presented for 2 sec. each at a 3.44-sec. rate with a 6.88-sec. intertrial interval, *S* was required to learn an invariant sequence of four pushbutton responses. The criterion of mastery was two successive error-free trials. Response availability was manipulated by means of a set of detachable covers for the apparatus response keys.

It was found that N_R exerted a highly significant retarding effect upon the rate of trial-and-error learning, in accordance with expectation.

Probability of correct first choice (p_+) for three representative values of N_R yielded ogival acquisition curves as functions of trials. The curves originated at hypothetical p_+ values consistent with the initial chance solubility of each problem. The apparent inflection points were a direct function of N_R while the acquisition rates varied inversely.

References

1. Friedman, M. Use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Amer. statist. Ass.*, 1937, **32**, 675–701.
2. Hull, C. L. *A behavior system.* New Haven: Yale Univ. Press, 1952.
3. McGeoch, J. A. *The psychology of human learning.* New York: Longmans, Green, 1942.
4. Melton, A. W. Learning. In W. S. Monroe (Ed.), *Encyclopedia of educational research.* (Rev. Ed.) New York: Macmillan, 1950. Pp. 668–690.
5. Noble, C. E. An analysis of meaning. *Psychol. Rev.*, 1952, **59**, 421–430.
6. Noble, C. E. The role of stimulus meaning (*m*) in serial verbal learning. *J. exp. Psychol.*, 1952, **43**, 437–446.
7. Noble, C. E. The meaning-familiarity relationship. *Psychol. Rev.*, 1953, **60**, 89–98.
8. Noble, C. E., & Farese, F. J. An apparatus for research in human selective learning. *J. Psychol.*, 1955, in press.
9. Riley, D. A. Rote learning as a function of distribution of practice and the complexity of the situation. *J. exp. Psychol.*, 1952, **43**, 88–95.
10. Spence, K. W. Theoretical interpretations of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 690–729.
11. Underwood, B. J. *Experimental psychology.* New York: Appleton-Century-Crofts, 1949.

(Received April 14, 1954)

Journal of Experimental Psychology
Vol. 49, No. 2, 1955

A TEST OF A MODEL FOR MULTIPLE-CHOICE BEHAVIOR [1]

MARVIN H. DETAMBEL

University of Southern California and U. S. Navy Electronics Laboratory[2]

This paper is concerned with experiments in which *S*s are required to predict on each of t trials which event of n possible events will occur; it deals specifically with experiments in which the stimulus that initiates a prediction contains no clue as to which event will occur (**3, 8, 9, 10, 11, 12**). A mathematical model, outlined below, has been constructed to describe behavior in repetitive multiple-choice situations of this sort. In constructing this model the author has been greatly influenced by the prior work of Estes (**5**), Estes and Burke (**6**), and Bush and Mosteller (**1, 2**).

An experimental test of some of the model's predictions is reported.

The Model

We deal with the case of any fixed number, n, of response alternatives, A, B, C, D, . . . , n, where each response alternative is a prediction that a particular one of n possible events will occur. The dependent variable of the model is pA_t, the probability that on trial t S will make prediction A.

We assume that pA changes on each trial, the amount of change, ΔpA_t, being determined by a mathematical operation upon pA_t. This mathematical operation is a joint function of a number of factors, of which the following are represented in the model:

1. Information given to S on trial t,
2. The number of response alternatives (possible predictions),
3. θ, a parameter which varies in value among *S*s, but is assumed constant for a given experiment. θ is the equivalent of $\bar{\theta}$ in Estes and Burke (**6**).

In mathematical notation,

$$pA_{t+1} = pA_t + \Delta pA_t \qquad (1)$$

where

$$\Delta pA_t = f(pA_t). \qquad (2)$$

Proceeding toward a statement of the particular function which ΔpA_t is of pA_t, we shall consider two classes of experimental operations for giving S information about the events being predicted:

Contingent class.—Under this class of operations the information given to S on any trial is dependent upon the prediction he makes on that trial. If he predicts the event which in fact occurs on that trial, he is informed that he has predicted correctly; if he predicts any other event, he is informed that his prediction was incorrect.

Noncontingent class.—Under this class of operations the information given on any trial is independent of the prediction S makes. Following each prediction S is informed which event actually occurred on that trial, i.e., which prediction was correct on that trial.

In the two sections which follow specific relationships are postulated between these informational procedures and response probability, and a number of predictions are derived concerning *S*s' performance in multiple-choice experiments.

Noncontingent Procedure

As indicated above, at the conclusion of each trial in the noncontingent procedure, S is informed which prediction was correct on that trial. As others have done, we assume that the information that a particular prediction was correct increases the likelihood of future occurrence of that prediction. Specifically, the increase in probability is assumed to be proportional to the

[1] This report covers work on the University of Southern California–U. S. Navy Electronics Laboratory Joint Project No. 6. It is part of a series of investigations under Contract No. NObsr-57435 between the Bureau of Ships, Navy Department, and the University of Southern California in cooperation with the U. S. Navy Electronics Laboratory. The experimental work was carried on at the U. S. Navy Electronics Laboratory.

[2] Now at the University of Illinois.

maximum possible increase. Thus, information that A was the correct prediction on trial t results in

$$\Delta pA_t = \theta(1 - pA_t). \tag{3}$$

Similarly, when at the conclusion of a trial S is informed that a *not-A* prediction, e.g., prediction B, was correct, there is assumed to be an increase in pB proportional to its maximum possible increase. Since the probabilities of the n possible predictions must total 1.0 at all times, any change in the probability of one must be accompanied by a change in the probability of at least one other prediction. We assume specifically that when an experimental operation increases the probability of one of the predictions there is a compensatory decrease in the probability of *each* of the other predictions, the amount of probability decrement for each alternative being proportional to the maximum possible decrement. Thus, the change in pA resulting from information that B was the correct prediction on trial t is

$$\Delta pA_t = -\theta pA_t. \tag{4}$$

In a random reinforcement situation, if we let $\pi A, \pi B, \pi C, \ldots, \pi n$ represent the proportion of trials on which A, B, $C, \ldots, n$ are correct, Equation 3 will be applicable on πA of the trials and Equation 4 on $(\pi B + \pi C + \ldots + \pi n)$ or $(1 - \pi A)$ of the trials. Then the average or expected change in probability of A on trial t is given by the relation

$$\begin{aligned} E\Delta pA_t &= \theta(1 - pA_t)(\pi A) \\ &\quad - \theta(pA_t)(1 - \pi A) \\ &= \theta(\pi A - pA_t). \end{aligned} \tag{5}$$

Now, substituting Equation 5 for ΔpA_t in Equation 1, the expected probability of A on any trial can be expressed, given its probability on the immediately preceding trial. If the probability on the first trial of an experiment is pA_1, then on the second trial

$$\begin{aligned} EpA_2 &= pA_1 + \theta(\pi A - pA_1) \\ &= \theta(\pi A) + pA_1(1 - \theta), \end{aligned}$$

on the third trial

$$\begin{aligned} EpA_3 &= \theta(\pi A) + pA_1(1 - \theta) \\ &\quad + \theta\{\pi A - [\theta(\pi A) + pA_1(1-\theta)]\} \\ &= \theta(\pi A) + (1 - \theta) \\ &\quad \times [\theta(\pi A) + pA_1(1 - \theta)] \\ &= \pi A - (\pi A - pA_1)(1 - \theta)^2, \end{aligned}$$

and so on for successive trials. In general, it can be shown that on trial $(t + 1)$

$$\begin{aligned} EpA_{t+1} &= \pi A \\ &\quad - (\pi A - pA_1)(1 - \theta)^t. \end{aligned} \tag{6}$$

Since θ is a quantity between 0 and 1, Equation 6 is seen to be a negatively accelerated curve beginning at pA_1 and approaching πA asymptotically.

An expression equivalent to Equation 6 has been presented by Estes and Straughan (7) for the special case of $n = 2$. If the assumptions underlying the present model are correct, the same equation applies for all n, $n \geqslant 2$, and given a constant πA and θ, the same asymptotic value will be approached at the same rate regardless of the number of alternative predictions available to S.

Contingent Procedures

We shall consider only the class of contingent experiments in which one of the predictions A is correct on some proportion πA of its occurrences and incorrect on the remaining $(1 - \pi A)$, and all other predictions are always incorrect.

When S predicts A and is informed that his prediction was correct, the conditions have been met under which Equation 3 was said to be applicable. Therefore, since $(pA_t)(\pi A)$ is the joint probability of A occurring and being correct on trial t, the contribution of correct A predictions to the average or expected probability change on any trial is $\theta(1 - pA_t)(pA_t)(\pi A)$.

On those trials when S predicts A and is informed that his prediction was incorrect, he is not told which of the remaining $n - 1$ predictions was correct.[3] Such information is assumed to reduce

[3] However, in the case of two alternatives, one of which must be correct, the information "A was incorrect" is logically equivalent to the information "B was correct," and both pieces of information may be expected to have equivalent effects upon pA. Consideration of this special case led to use of the same constant of proportionality, θ, in the expressions for probability increments (Equation 3) and decrements (Equation 8).

the probability of A by an amount proportional to the maximum possible decrease. Thus, information that A was an incorrect prediction on trial t results in

$$\Delta p A_t = -\theta p A_t. \tag{8}$$

Since the probability of incorrectly predicting A on any trial is $(pA_t)(1-\pi A)$, the contribution of such events to the average probability change on any trial is $(-\theta pA_t)(pA_t)(1-\pi A)$.

Whenever prediction of a *not-A* alternative occurs, S is informed that he has predicted incorrectly, resulting in reduced probability of that *not-A* alternative. Again, in order for the n probability values to total 1.0, there must be a compensatory increase in the probability of at least one alternative. It is assumed that the compensatory increment is distributed among *all* of the other predictions in proportion to the maximum possible increment of each. Therefore, the change in pA resulting from information that B was an incorrect prediction on trial t is expressed by the relation

$$\Delta p A_t = \theta p B_t \times \frac{(1-pA_t)}{(1-pA_t)+(1-pC_t)+(1-pD_t)+\ldots+(1-pn_t)}$$

$$\theta p B_t = \frac{1-pA_t}{n-2+pB_t}, \tag{9}$$

or substituting as an approximation for pB_t the mean probability of *not-A* alternatives, $\frac{1-pA_t}{n-1}$, we have

$$\Delta p A_t = \theta\left(\frac{1-pA_t}{n-1}\right)\times\left(\frac{1-pA_t}{n-2+\frac{1-pA_t}{n-1}}\right) = \theta\frac{(1-pA_t)^2}{(n-1)(n-2)+1-pA_t}. \tag{10}$$

Equation 10 multiplied by the probability of *not-A* alternatives $(1-pA_t)$ is the contribution to expected probability change that results from incorrect predictions of *not-A* alternatives.

Now, by combining the contributions to $E\Delta pA_t$ made by (a) correct predictions of A, (b) incorrect predictions of A, and (c) incorrect predictions of *not-A* alternatives, we have

$$\begin{aligned} E\Delta pA_t &= \theta(1-pA_t)(pA_t)(\pi A) - \theta(pA_t)(pA_t)(1-\pi A) \\ &\quad +\theta(1-pA_t)\left[\frac{(1-pA_t)^2}{(n-1)(n-2)+(1-pA_t)}\right] \\ &= \theta\left[(pA_t)(\pi A-pA_t) + \frac{(1-pA_t)^2}{(n-1)(n-2)+(1-pA_t)}\right]. \end{aligned} \tag{11}$$

By substituting Equation 11 in Equation 1 and following the procedure used in deriving Equation 6 it can be shown that in the special case of $n = 2$

$$EpA_{t+1} = \frac{1}{2-\pi A} - \left(\frac{1}{2-\pi A} - pA_1\right)[1-\theta(2-\pi A)]^t. \tag{12}$$

When $\theta(2-\pi A) > 1$, this function oscillates about the value $\frac{1}{2-\pi A}$. However, experimental evidence indicates that values of θ are sufficiently small so that $\theta(2-\pi A) < 1$. Under these circumstances Equation 12 is a negatively accelerated curve beginning at pA_1 and approaching $\frac{1}{2-\pi A}$ asymptotically.

For $\pi A = 1$ Equation 12 reduces to Equation 6, and both the contingent and noncontingent curves approach an asymptote of 1.0 at rates determined solely by θ. For all other values of πA, with fixed θ, $\theta(2-\pi A) > \theta$, i.e., the learning rate parameter is always greater in the two-alternative contingent case than in the noncontingent case.

The asymptotes of Equations 6 and 12 have already been given as πA and $\frac{1}{2-\pi A}$. It is easy to show that $0 \leqslant \pi A < 1$ implies $\pi A < \frac{1}{2-\pi A}$, and therefore it is a characteristic of the model that for fixed πA the asymptote is always lower in the noncontingent case than in the two-alternative contingent case. The asymptotes of the two curves will be the same when πA (noncontingent)

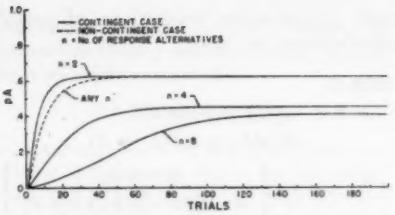

FIG. 1. Illustrative curves showing probability of A response as a function of number of trials for experimental conditions described in the text. For all curves $\theta = .1$, and initial probability is 0.

$= \dfrac{1}{2 - \pi A\ \text{(contingent)}}$. Two curves having this relationship are shown in Fig. 1.

For $n \geq 3$ no statement of Ep_{A_t} has been found that is comparable in brevity and computational ease to Equations 6 and 12. However, for any n, probability curves can be generated by computing a series of solutions to Equation 1 for trials 1, 2, 3, . . . , t.

Several illustrative curves are shown in Fig. 1. The curves for the contingent case were computed using Equation 11 as the statement of $E\Delta p_{A_t}$, with $\pi A = .4$, and separate computations for $n = 2$, 4, and 8. In the curve for the noncontingent case $\pi A = .625$. In all four curves $\theta = .1$ and $p_{A_1} = 0$.

If we think of Ep_{A_t} as the expected *frequency* of A on trial t, then $\sum_{\tau=1}^{t} Ep_{A_t}$ is the number of A responses that will be expected to occur in t trials. Such summations over trials were computed for the curves in Fig. 1, resulting in the cumulative response curves of Fig. 2. Curves of this sort are particularly useful when a comparison is to be made between model-predicted performance and the performance of one S or a small group of Ss.

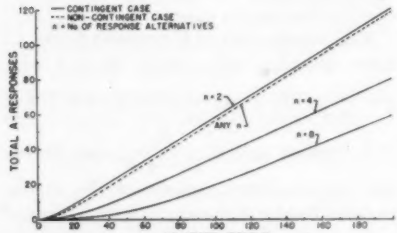

FIG. 2. Illustrative cumulative frequency curves, produced by summing, over trials, the probability functions in Fig. 1.

In the contingent curves in Fig. 1 it will be noted that, with constant πA and θ, both the asymptotes and rates of change decrease with increasing n.

Also, as Fig. 1 suggests, the predicted curve becomes sigmoid as n increases. The precise conditions under which the S shape should appear in experimental data have not been investigated.

By setting Equation 11 equal to zero and solving, it can be verified that, with πA constant, the asymptotic value decreases as the number of alternatives becomes larger. The asymptote is $\dfrac{1}{2 - \pi A}$ for $n = 2$, and decreases, approximating πA more and more closely as n is increased.

In a similar way, Equation 11 can be used to demonstrate that with constant asymptote and θ, πA is an increasing function of n, and the learning rate parameter a decreasing function of n.

THE EXPERIMENTS

One of the reported experiments utilized the values of n and πA used in constructing the curves in Fig. 1 and 2. Except for the choice of θ and p_{A_1}, which may vary from S to S, these curves should describe the results of the experiment if the model is to be considered descriptively adequate.

The second experiment was an attempt to gain some information about the parameter θ. In the first experiment in which S participates θ is a curve-fitting constant. It is possible, however, that after the value of θ has been determined for an individual in one experiment, this value can be used in predicting his performance in other experiments. This will be true if θ remains constant or if it changes in some known way from experiment to experiment. On the assumption that we should be able to evaluate θ from the data of the

experiments outlined above, all *Ss* were run in a second experiment in an effort to determine what change in θ, if any, would occur.

Method

Forty-eight students at San Diego State College were paid to participate in two short experiments. Total time required of each *S* was about 45 min.

The *E* and *S* sat on opposite sides of a table with a large opaque screen separating *E* and his apparatus from *S*. The *S* was seated before a metal box mounted so that the face of the box slanted away from him at a 45° angle.

The box contained a horizontal row of eight push buttons at 1¾-in. intervals, and 3½ in. above the buttons, a parallel row of eight small electric lights. Each light was directly above a button and each button-light pair was joined by a heavy line so there would be no mistaking which light was paired with which button. All eight pairs of lights and buttons could be exposed, or cover plates could be bolted to the response box so as to expose only two or four pairs.

In both experiments *S*'s task was to predict on each of 200 trials which one of the *n* lights could be turned on; the prediction was made by pushing the button paired with the predicted light. A buzzer, activated for .5 sec. at regularly spaced 4.5-sec. intervals, was the signal for *S* to indicate his prediction.

In Exp. I, *Ss* in the Noncontingent group (NC) were told that one or another of the lights would be turned on each time they pushed a button, and that they were to try always to push the button paired with the light that would be turned on. The lighting of a light following each prediction informed *S* which button he should have pushed on that trial.

In Exp. I, *Ss* in the Contingent group (C) were told that on each trial one or another of the lights could be turned on, that they should predict which light could be turned on and indicate their prediction by pushing the button paired with the predicted light. Here a correct prediction was indicated if a light was turned on; incorrect predictions never turned on a light. Under this procedure *S* was informed whether his prediction was correct or incorrect, but if incorrect he was not informed which prediction was the correct one to have made on that trial.

The *E* secretly designated one of the buttons the "*A* alternative." Under every condition of the first experiment each of the available buttons was *A* for an equal number of *Ss*.

In the NC group the *A* light was turned on following a randomly determined 62% of the 200 responses; this is as close an approximation as the apparatus permitted to the $\pi A = .625$ which is a parameter of the broken curves in Fig. 1 and 2. Each of the *not-A* lights lighted about equally often on the remaining 38% of the 200 trials.

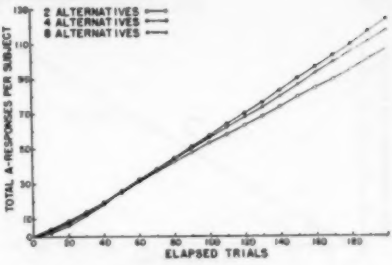

FIG. 3. Cumulative *A*-response curves for the Noncontingent groups, Exp. I.

In the C group the *A* light was turned on by 40% of the predictions of *A*; this corresponded to the parameter $\pi A = .4$ in the solid curves of Fig. 1 and 2. Under this procedure the *not-A* lights never lighted.

The 24 NC *Ss* were equally distributed among the NC-2, NC-4, and NC-8 subgroups. The numeral indicates the number of button-light pairs available to *S*. Similarly, the 24 C *Ss* were equally distributed among the C-2, C-4, and C-8 subgroups.

After 200 trials *S* rested about 5 min. before beginning Exp. II. During this interval *E* prepared his apparatus for the contingent, four-alternative procedure (C-4). Then, after appropriate instructions, all *Ss* were given 200 trials under the C-4 procedure. In this experiment the *A* button was the third button from the left for all *Ss* except eight for whom *A* was the second button from the left. Four of these eight *Ss* were in the C-2 group in Exp. I, and four were in the NC-2 group.

All *Ss*' predictions were recorded automatically on a moving paper tape.

Results

Experiment I

Noncontingent procedure.—Figure 3 consists of cumulative frequency curves for the three NC groups. The plotted points represent the number of *A* responses per *S* that occurred during the first 10, 20, 30, . . . , 190, 200 trials of Exp. I.

An analysis of variance for repeated measurements on the same *Ss* was performed, using as a response measure the number of *A* responses within successive 20-trial blocks (**4,** p. 288).

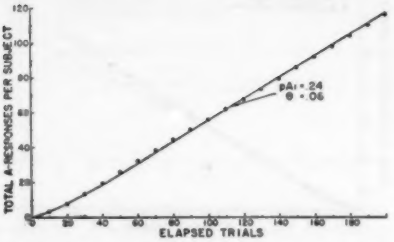

FIG. 4. Cumulative *A*-response data for the combined Noncontingent groups, with a best fitting curve generated by the model.

This analysis shows statistically significant differences due to trials ($F = 8.63$, 9 and 189 *df*, $p < .01$). Not statistically significant are the small differences among groups in total *A* responses ($F = 2.13$, 2 and 21 *df*), and differences due to interaction between trials and treatments ($F = 1.18$, 18 and 189 *df*). In the absence of evidence that the groups differed systematically, the data from the three groups were combined into the cumulative response curve of Fig. 4. The circles represent observed frequencies. The solid curve was generated by the model, using $\theta = .06$ and $pA_1 = .24$.

Contingent procedure.—Cumulative *A*-response data for the three C groups are presented in Fig. 5. Significant differences in frequency of *A* responses are attributable to trials ($F = 12.29$, 9 and 189 *df*, $p < .01$), treatments ($F = 6.22$, 2 and 21 *df*, $p < .01$), and interaction between trials and treatments ($F = 1.52$, 18 and 189 *df*, $p < .05$).

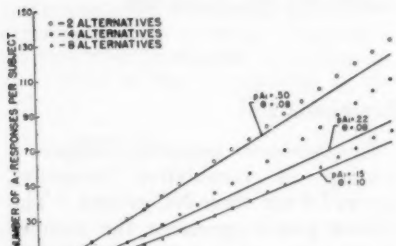

FIG. 5. Cumulative *A*-response data for the Contingent groups, Exp. I. Model-generated curves are fitted to the data obtained during the early trials.

The model-generated curves in Fig. 5 are good representations of the data over at least the first 80 trials, but in later trials they underestimate the rate of occurrence of *A*. The difference between predicted asymptotic frequency of *A* responses and the observed frequency during the final 30 trials was significant for the C-4 group ($t = 2.60$, 7 *df*, $p < .05$) but not for the C-2 and C-8 groups. On this evidence, and assuming that the groups had attained a stable response rate, we could argue that the model adequately predicted asymptotic frequency of *A* responses for two of the groups. This is not a very convincing argument however, and it is further weakened by the fact that the groups did not differ significantly from one another in frequency of *A* responses during these 30 trials ($F = 1.68$, 2 and 21 *df*). These facts suggest that during these 30 trials within-groups variance was sufficiently large to prevent rejection of a variety of null hypotheses.

Under informal questioning at the conclusion of the experiment many *S*s stated their belief that in the procedure which allowed only correct predictions to turn on lights, only one of the buttons was ever correct. Many *S*s said that for this reason they selected this one button on most trials, only occasionally pushing another to be sure the situation had not changed. This decision that only one alternative was ever correct seems to manifest itself rather suddenly in the data of a number of *S*s. For example, in Fig. 6 are plotted the cumulative

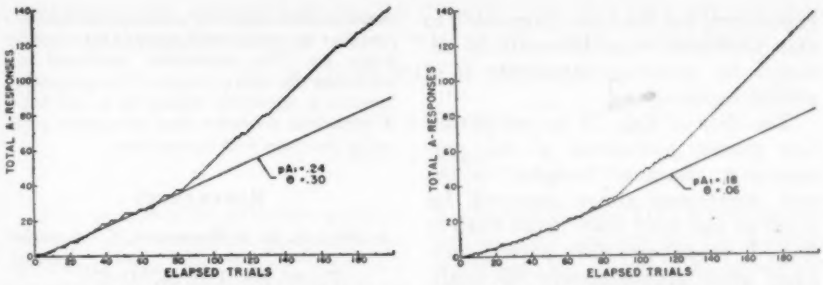

FIG. 6. Cumulative A-response curves for two Ss under the C-4 procedure, showing abrupt increases in frequency of A response at about Trial 84. The solid lines are model-generated curves of best fit.

response curves of two Ss. The solid curves generated by the model show that Ss were responding in agreement with the model for some 80 trials, and then abruptly increased their relative frequency of A responses. Such "insights" are at least partially responsible for the greater than predicted rate of A responses during the later trials of both experiments. To avoid misrepresentation it should be pointed out that the curves of many Ss did not fit the model as exactly as the data plotted in the early trials of these two examples.

Experiment II

Cumulative A-response curves for Exp. II are presented in Fig. 7. Significant differences in frequency of A responses are attributable to trials ($F = 37.87$, 9 and 476 df, $p < .01$), and Trial × Treatment interaction ($F = 1.92$, 45 and 476 df, $p < .01$). Differences due to treatments were not statistically significant ($F < 1.00$).

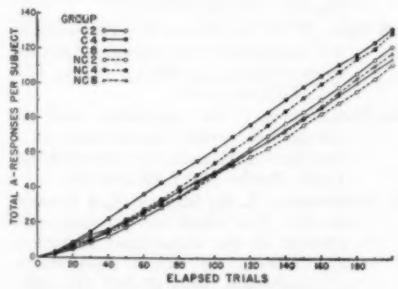

FIG. 7. Cumulative A-response curves, all groups, Exp. II.

Since all groups were run under identical conditions in Exp. II, the significant Trial × Treatment interaction is evidence of differential transfer effects produced by the several procedures employed in Exp. I.

While the procedures of Exp. I differentially affected the rate of growth of the curves in Exp. II, they did not produce differences in the final relative frequency of A. Differences in number of A responses during the final 30 trials are not statistically significant ($F < 1.0$). Further, no group's performance differed significantly from the C-4 group's frequency of A during the final 30 trials of Exp. I. This was established by a t test of the largest such difference ($t = 1.07$), and since this difference was not statistically significant, it was presumed that the smaller differences would also be nonsignificant.

In summary, when the C-4 procedure was used, the same final rate of A occurrences was approached regardless of the amount or type of previous

experience, but the rate of approach to this final level was differentially affected by previous experience in a similar experiment.

The data of Exp. II do not permit any precise evaluation of the parameter θ. Because "insights" of the sort mentioned above occurred for more *S*s and after fewer trials than in Exp. I, model and data parted company after approximately 40 trials. Even though the model can provide curves that fit the data over the first 40 trials, the fact that some of the abrupt shifts to a high probability level occurred within these early trials suggests that the resulting θ's would be spuriously high.

Available data do not indicate how the model should be modified to take into account the suddenly increased rate of A responses exhibited by some *S*s under the contingent procedure. Studies relating experiment and subject parameters to occurrence or nonoccurrence of sudden probability changes, and to the date and extent of such changes, may lead to useful modifications of the model.

Summary

A mathematical model was outlined for experiments in which *S*s are required to predict on each of t trials which event of n possible events will occur. An experimental test of some of the model's predictions was described.

Excellent agreement between data and model was found when *S*s were informed, following each of their predictions, which one of the n possible events actually did occur. Here, the probability that S will predict a particular one of the events is a negatively accelerated growth or decay function of trials, approaching as a limit the relative frequency of the event being predicted.

When *S*s were informed only that their prediction was correct or incorrect, model-generated curves fitted the response data over the early trials of the experiment, but failed to do so in later trials when some *S*s abruptly increased their frequency of prediction of one of the alternatives. Under these conditions the model predicts negatively accelerated probability functions when n is small, sigmoid functions for larger n's. The experiment confirmed the prediction that rate of change of the probability function is negatively related to n, but failed to verify the prediction that asymptotic probability decreases with increasing n.

References

1. Bush, R. R., & Mosteller, F. A mathematical model for simple learning. *Psychol. Rev.*, 1951, **58**, 313–323.
2. Bush, R. R., & Mosteller, F. A model for stimulus generalization and discrimination. *Psychol. Rev.*, 1951, **58**, 413–423.
3. Detambel, M. H. A reanalysis of Humphreys' "acquisition and extinction of verbal expectancies." Unpublished master's thesis, Indiana Univ., 1950.
4. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
5. Estes, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, **57**, 94–107.
6. Estes, W. K., & Burke, C. J. A theory of stimulus variability in learning. *Psychol. Rev.*, 1953, **60**, 276–286.
7. Estes, W. K., & Straughan, J. H. Analysis of a verbal conditioning situation in terms of statistical learning theory. *J. exp. Psychol.*, 1954, **47**, 225–234.
8. Grant, D. A., Hake, H. W., & Hornseth, J. P. Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. *J. exp. Psychol.*, 1951, **42**, 1–5.
9. Grant, D. A., Hornseth, J. P., & Hake, H. W. The influence of inter-trial interval on the Humphreys' "random reinforcement" effect during the extinction of a verbal response. *J. exp. Psychol.*, 1950, **40**, 609–612.
10. Hake, H. W., & Hyman, R. Perception of the statistical structure of a random series of binary symbols. *J. exp. Psychol.*, 1953, **45**, 64–74.
11. Humphreys, L. G. Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *J. exp. Psychol.*, 1939, **25**, 294–301.
12. Humphreys, L. G., Miller, J., & Ellson, D. G. The effect of the inter-trial interval on the acquisition, extinction, and recovery of verbal expectations. *J. exp. Psychol.*, 1940, **27**, 195–202.

(Received March 16, 1954)

Journal of Experimental Psychology
Vol. 49, No. 2, 1955

STIMULUS GENERALIZATION OF AN INSTRUMENTAL RESPONSE AS A FUNCTION OF THE NUMBER OF REINFORCED TRIALS[1]

GARRY MARGOLIUS

Boston University

The existence of stimulus generalization and stimulus generalization gradients has had ample empirical verification since Pavlov (**13**) and Anrep (**1**) first reported data relevant to the concepts (**2, 3, 4, 6, 8, 9**). Despite the abundance of such demonstrations there is little information available pertaining to the many possible parameters of generalization gradients. It is felt that if data were made available as to possible changes in the generalization gradient following the manipulation of one or more variables in the experimental situation then predictions based upon the concept would obtain a higher degree of precision (**11, 12, 16**).

A positive relationship between absolute generalization and the number of conditioning trials has been reported by Hovland (**7, 10**). Relative generalization, however, was found to increase during the initial conditioning trials but progressively declined with additional training. Unfortunately, as only one of the many possible generalization test points was investigated we have no available data on *gradient* changes as a function of training trials in the classical conditioning situation. The findings of Razran (**15**), based upon a survey and analysis of the Russian literature, are in essential agreement with those reported by Hovland. Once again, the absence of data relevant to gradient changes limits the generality of the findings.

The present study was designed to enable observations of changes in the generalization gradient as the number of reinforced training trials is increased. More specifically, the study was viewed as an attempt to investigate the relationship existing between generalization gradients and training trials when the stimulus dimension is area, the *S* an albino rat, the method of training is instrumental conditioning, and the degree of generalization is measured by several possible criteria. It was felt that if rats were first trained to make a specific response to a training stimulus, and then after varying amounts of such training each rat was tested at one of several possible points on the stimulus continuum, several generalization gradients would be made available for analysis.

METHOD

Subjects.—The *S*s were 163 experimentally naive male albino rats from the colony maintained by the psychology department of the State College of Washington. The ages ranged from 92 to 126 days at the beginning of the experiment.

Apparatus.—The apparatus, similar to that described by Grice (**5**), was a 24-in. runway, 5 in. wide, 6 in. high, and covered with glass. It contained vertically sliding doors at each end, and one in the center. The runway was mounted

[1] This article is based on a dissertation submitted to the Graduate School of the State College of Washington in partial fulfillment of the requirements for the Ph.D. degree. The author wishes to express his appreciation to Dr. David Ehrenfreund for his advice and cooperation during the course of this research. A summary of this paper was read before the American Psychological Association meeting in Cleveland, 1953.

on a centered pivot so that either end could become in turn the start sector, and the response sector. A detachable 2-in. continuation of the runway contained the stimulus in use. The stimuli were white circles of 79, 63, 50, 32, and 20 sq.cm. in area, cut from 3/32-in. aluminum. Each circle was fastened to a panel by means of a 2-in. bracket. When a panel was inserted in the runway continuation the stimulus was flush with the end of the runway and equidistant from the floor, top, and side walls. A centered, counterweighted, 1 sq.-in. door set flush with the circle surface opened directly on a food cup. Two 10-w. incandescent bulbs provided all the light in the experimental room. They were placed 8 in. from the end of each sector of the runway, and both lights were 14 in. above the floor of the runway. Trial latencies were recorded on a Standard Electric Time clock.

Preliminary training.—Ten days before training all *S*s were placed on a food maintenance schedule. This consisted of a 1-hr. daily feeding period which followed the training session by 1 hr. The preliminary training occupied four days, during which all *S*s were treated alike.

On Days 1 and 2 *S* was taken to the experimental room at the time of regular feeding and placed in an 8 in. long, 4 in. wide, 5 in. deep black box which was illuminated by a 10-w. incandescent bulb placed 14 in. above the floor. Food pellets, similar to those used in training, were scattered about the floor, and *S* was allowed to eat freely for 20 min. The *S* was then returned to an individual feeding stall in the animal room, and allowed to eat for 40 min.

On Day 3 *S* was placed in the response end of the runway with the center door lowered. The food door in the 79 sq. cm. stimulus circle was open, and *S* was allowed to seize and eat a 100-mg. pellet placed in the food cup. After *S* had seized the pellet the vertical door in front of the stimulus was lowered and another pellet was placed in the food cup. When *S* had consumed ten pellets he was removed for 15 min. The runway was then rotated, and *S* was returned to the apparatus and allowed to remove and eat ten additional pellets from the open food door. The vertical door in front of the stimulus was lowered for approximately 1 min. after each pellet was secured.

On Day 4, after the first pellet was secured, the runway was rotated with *S* in the apparatus. The center door was then raised and *S* walked the length of the runway to obtain the second pellet. Following removal of the pellet from the food cup, both the door in front of the stimulus and the center door were lowered. For each succeeding pellet the runway was rotated, and for each succeeding pellet the food door in the stimulus circle more closely approximated the closed position. For Pellets 9 and 10 the door was almost completely closed. Nine *S*s failed to complete pretraining and were discarded.

Reinforcement training.—The first two instrumental training trials began as a continuation of preliminary training on Day 4. A trial was run in the following manner. The *S* was placed in the starting compartment with the center door lowered. As soon as *S* faced the center door, the door was raised, starting a chronoscope. When *S* pushed in the closed food door of the 79 sq.cm. circle, stopping the chronoscope, both the door in front of the stimulus and the center door were lowered and the runway was rotated. After *S* consumed the obtained pellet and was once again facing the center door, the next trial was given. All reward pellets used during training weighed approximately 200 mg.

The *S*s that completed preliminary training were randomly assigned to one of six groups. The groups differed only in terms of the number of reinforced trials to the 79 sq.cm. circle that they were to receive. The number of *S*s in the 0, 4, 16, 64, 104, and 164 reinforced trial conditions was 7, 35, 35, 35, 35, and 7, respectively. The spacing of training trials following the first two reinforcements on Day 4 was as follows. The group that received four reinforced trials received Trials 3 and 4 on Day 5 which was also their generalization test day. The 16-trial group received 12 trials on Day 5, and Trials 15 and 16 on Day 6, their generalization test day. The 64-, 104-, and 164-trial groups received 20 trials a day for the necessary number of days, and then received two additional trials on the generalization test day. Thus all training groups received their first two trials as a continuation of pretraining on Day 4, and their last two training trials on their test day. The group that received no reinforced trials following preliminary training received no reinforced trials on its test day.

Generalization tests.—The *S*s in reinforcement groups 4, 16, 64, and 104 were divided into five test groups corresponding to the 79, 63, 50, 32, and 20 sq.cm. circles. After the final two reinforced responses to the 79 sq.cm. circle *S* was tested for generalized response strength by 30 nonrewarded or extinction trials to the stimulus being investigated. The *S*s in both the 0-trials group and the 164-trials group were tested only on the training stimulus. This made available for later analysis performance at both extremes of the reinforcement schedule. Test trials were similar to training trials except that during the extinction trials the food door was locked so that it opened only ¼ in. and the food cup was empty. For *S*s being tested on

the 79 sq.cm. circle a new identical stimulus was substituted. If *S* failed to respond within 60 sec. the trial was scored as a failure of response. The apparatus and procedure utilized during the test series made available for analysis three response criteria.

The response criteria were: (*a*) Latency: As the chronoscope was started when *E* raised the center door and as it was stopped when *S* pushed in the food door, a latency measure was available for each test trial. Each latency measure was expressed in terms of its reciprocal to facilitate comparison with the other criteria. (*b*) Total number of responses: As each *S* was given 30 opportunities to respond to the test stimulus, a simple counting of the number of responses out of 30 was considered as one measure of response strength. (*c*) Number of responses in 60 sec.: After each trial the chronoscope was reset to zero. However, by adding each succeeding latency in a cumulative fashion until the total of 60 sec. was reached, a criterion expressing the number of responses within a specified time limit was made available.

Results

Absolute generalization.—The reciprocal of the median latency for the first three test trials × 1000 was computed for all stimuli tested at all levels of training. The data are presented graphically in Fig. 1. Here the latency index is plotted as a function of area of the test stimuli in logarithmic units. For this criterion the gradients are negatively accelerated for all training levels. Further, all gradients indicate a positive relationship between absolute generalization and training. It may be noted that a plot of the first test trial latencies resulted in gradients that are in essential agreement with the data presented here.

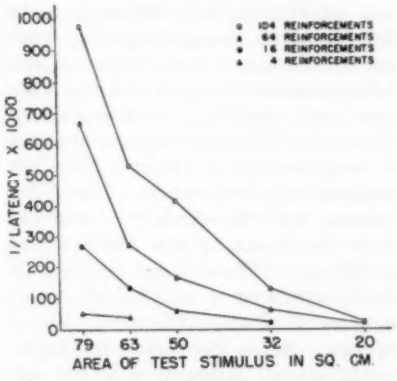

FIG. 1. Generalization gradients based on the median latency of the first three test trials as a function of stimulus area and training trials. The 0- and 164-trial training groups have not been included.

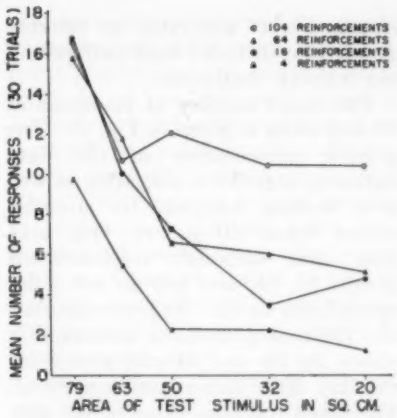

FIG. 2. Generalization gradients based on responses in 30 test trials as a function of stimulus area and training trials. The 0- and 164-trial training groups have not been included.

Test responses of the 0- and 164-trial groups were obtained so that an evaluation of response strength to the training stimulus at extreme values of the reinforcement schedule might be made. The zero training group performance score of 38 is considered as the additive combination of chance responses, plus whatever response strength that may have developed during preliminary training. Comparison of the 104- and 164-trial values, 980 and 934 respectively, indicates that peak performance in terms of the latency criterion has been achieved by 104 trials. The behavior recorded indicates that the gradients

obtained were generated by selected points that included both extremes of the training continuum.

The mean number of responses in 30 test trials is given in Fig. 2. The graphic presentation of the data indicates negative acceleration as well as a leveling tendency for stimulus values below 50 sq.cm. One may note that although reinforcement groups 16, 64, and 104 do not differ appreciably at the 79 sq.cm. stimulus, the 104-trial gradient is considerably above the 16- and 64-trial groups for the last three test stimuli employed. The mean response value of the zero group to the training stimulus in the test situation was 6.29. A comparison of the test results of the 104- and 164-trial training groups, 16.71 and 15.00 respectively, indicates that peak performance to the 79 sq.cm. stimulus has been achieved by 104 trials. In fact, for this criterion, the peak has been reached by 16 reinforcements.

The analysis of variance technique was applied to test the hypotheses that the means were equal for levels of training, and for the stimulus values under investigation. The *F* for training is 25.58, which for 3 and 132 *df* is significant at the .001 level of confidence. The *F* for stimuli is 29.19, which for 4 and 132 *df* is significant at the .001 level of confidence. As the differences cannot be reasonably attributed to random sampling error the null hypothesis may be rejected in both cases for this criterion. The *F* ratio for training × stimuli interaction is not significant.

The mean number of responses made during the first 60 sec. of test trial time is presented in Fig. 3. All gradients are negatively accelerated, and once again a positive relationship between training and absolute generalization is evident. The mean response value of the zero training group to the 79 sq.cm. stimulus in the test situation was 1.14. A comparison of the test scores of the 104 and 164 training groups, 7.71 and 7.00 respectively, indicates that peak performance to the training stimulus has been achieved by 104 training trials.

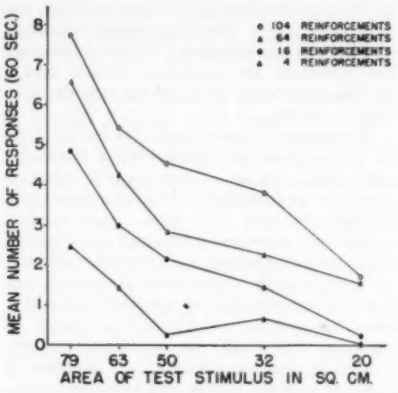

FIG. 3. Generalization gradients based on responses in 60 seconds as a function of stimulus area and training trials. The 0- and 164-trial training groups have not been included.

Application of Bartlett's chi-square test showed the cell variances to be heterogeneous, suggesting that analysis of variance was inappropriate. A log transformation, of the raw score plus one, yielded a distribution of values that fully met the requirement of homogeneity of variance, and the analysis was then made. The *F* for training is 45.00, which for 3 and 132 *df* is significant at the .001 level of confidence. The *F* for stimuli is 34.70, which for 4 and 132 *df* is also significant at the .001 level of confidence. Again, the *F* ratio for Training × Stimuli interaction is not significant. Thus the analysis for this criterion indicates the same statistical significance as that obtained by the total responses criterion. It

TABLE 1

RELATIVE GENERALIZATION (PERCENTAGE) COMPUTED FROM LATENCY MEASURES, RESPONSES IN 30 TEST TRIALS, AND RESPONSES IN 60 SECONDS

Training Trials	Stimulus in Square Centimeters			
	63	50	32	20
Latency Measures				
4	76	0	0	0
16	50	23	9	0
64	41	25	10	3
104	54	42	13	3
Responses in 30 Test Trials				
4	55	23	23	13
16	59	43	21	29
64	75	41	38	32
104	64	73	63	63
Responses in 60 Seconds				
4	59	12	29	0
16	62	44	29	6
64	65	44	35	24
104	70	59	50	22

would seem reasonable to assert that response strength to the generalized stimulus is inversely related to the distance of that stimulus from the training stimulus, and that as the number of reinforced responses to the training stimulus is increased the absolute amount of generalized response strength increases.

Relative generalization.—Table 1 presents the data pertinent to changes in relative generalization that accompany changes in the training schedule, for the three response criteria. Relative generalization is here considered as the ratio of generalized response strength to the response strength to the training stimulus. Such ratios were computed for each generalization stimulus with respect to the response score to the training stimulus for each level of training. For the latency measures, it may be noted that at the 63-sq.cm. test point relative generalization decreases with added training until 64 is reached, and then increases. For all other test stimuli an increase in relative generalization accompanies an increase in training trials. With respect to total number of responses in 30 test trials, and number of responses made within 60 sec. of test trial time, increased training resulted in a trend of increasing relative generalization for all stimuli tested.

To facilitate comparison of the present data with that reported by Hovland (**10**), response values for all generalization test points were combined, and a mean value computed for each training level. Relative generalization ratios computed from the values obtained are presented in Fig. 4. For the three criteria employed a positive relationship between training and relative generalization is evident.

DISCUSSION

Inspection of the results suggests that a meaningful discussion of generalization must take into account the performance measure employed. It should be noted that for the total number of responses in

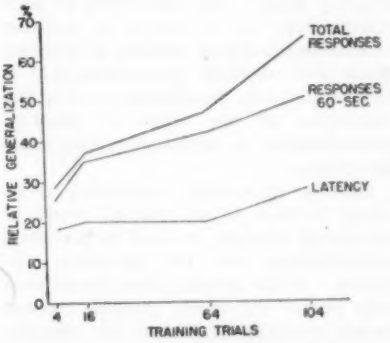

FIG. 4. Mean relative generalization of all test stimuli as a function of training trials and criterion.

30 trials criterion, peak performance to the training stimulus is reached after only 16 trials. For the other two criteria, this peak required 104 training trials. This seeming discrepancy emphasizes the necessity of considering both general procedure and response measure employed in any attempt to relate strength of response to number of training trials. Considering the possibility of the operation of effects such as inhibition, which would be reflected differentially in the various performance measures, it should not be too surprising to note that the three measures employed in the present investigation yield somewhat different constants for each level of training.

Utilizing the 30 trials criterion, the gradient obtained for the 64-trial training group appears to be in fair agreement with the gradient generated by 67 reinforcements reported by Grice and Saltz (**6**). The composite results obtained are further confirmation of the soundness of assuming decreasing generalization gradients. The gradients obtained are in striking agreement with Hull's assumption of a negatively accelerated theoretical gradient for generalization to areas smaller than the training stimulus. Although a rigorous fit was not feasible, inspection of the gradients obtained by any of the criteria indicates that both origin and asymptotic level are positively related to number of training trials. The net result of such relationships is, of course, a positive relationship between number of training trials and absolute generalization. It follows that any assumption relating the restriction of the range of absolute generalization to increased training must be rejected.

Another important relationship to be noted is that of an over-all trend of increasing relative, as well as absolute, generalization for the gradients obtained. It has already been shown that only one of the twelve possible column trends partially confirmed the assumption of decreasing relative generalization as the number of reinforced training trials is increased. With the exception of the relationship holding within the 63-sq.cm. column of the latency criterion (Table 1), an increase in relative generalization may be noted for all generalized stimuli. The combined results of all test stimuli at each level of training, Fig. 4, are clearly in agreement with the individual analyses. The data are not in agreement with the conclusions of Hovland (**10**) and the conclusions of Razran (**14, 15**). No evidence of "spontaneous differentiation" (**7**, p. 732) as measured by the decline of relative generalization with further reinforcements was found. In fact, an orderly increase of relative generalization with added training is clearly evident. It may be the case that classical techniques and instrumental techniques may yield contradictory data concerning relative generalization. This seems hardly likely when one considers the high degree of uniformity among acquisition and extinction curves obtained under either classical or instrumental conditioning procedures.

Summary

In a study designed to determine the relationship between number of training trials and the generalization gradient of a size-brightness dimension in an instrumental conditioning situation, four groups of 35 albino rats were trained to open a door in the center of a white circle, 79 sq.cm. in area. These groups received either 4, 16, 64, or 104 reinforced trials to the training stimulus.

After training each group was divided into five subgroups of seven *Ss* each which received 30 nonreinforced test trials to either the 79-sq.cm. training circle, or to a circle of 63, 50, 32, or 20 sq.cm. Two additional groups of seven *Ss* received 0 or 164 training trials, respectively, to the 79 sq.cm. stimulus, and both were then tested to that stimulus. Test performance was evaluated in terms of the total number of responses during the series, the number of responses made within 60 sec., and the median latency of the first three test trials.

The results were as follows:

1. Response strength to the generalized stimulus was found at all levels of training to be inversely related to the distance of that stimulus from the training stimulus.

2. All generalization gradients obtained were negatively accelerated when the response measure was plotted as a function of log area.

3. As the number of reinforced trials to the training stimulus was increased the height of the generalization gradient was found to increase.

4. As the number of reinforced trials to the training stimulus was increased both absolute and relative generalization were found to increase throughout the range investigated.

5. When both the training and stimulus dimensions were investigated there resulted a family of curves which is dependent upon both experimental variables and upon the criterion employed.

REFERENCES

1. ANREP, G. V. The irradiation of conditioned reflexes. *Proc. roy. Soc.*, 1923, **94B**, 404–425.
2. BASS, M. J., & HULL, C. L. The irradiation of a tactile conditioned reflex in man. *J. comp. Psychol.*, 1934, **17**, 47–65.
3. BROWN, J. S. The generalization of approach responses as a function of stimulus intensity and strength of motivation. *J. comp. Psychol.*, 1942, **33**, 209–226.
4. GIBSON, E. J. Sensory generalization with voluntary reactions. *J. exp. Psychol.*, 1939, **24**, 237–253.
5. GRICE, R. G. Visual discrimination learning with simultaneous and successive presentation of stimuli. *J. comp. Physiol. Psychol.*, 1949, **42**, 365–373.
6. GRICE, R. G., & SALTZ, E. The generalization of an instrumental response to stimuli varying in the size dimension. *J. exp. Psychol.*, 1950, **40**, 702–708.
7. HOVLAND, C. I. The effects of varying amounts of reinforcement upon the generalization of conditioned responses. *Psychol. Bull.*, 1935, **32**, 731–732.
8. HOVLAND, C. I. The generalization of conditioned responses: I. The sensory generalization of conditioned responses with varying frequencies of tone. *J. gen. Psychol.*, 1937, **17**, 125–148.
9. HOVLAND, C. I. The generalization of conditioned responses: II. The sensory generalization of conditioned responses with varying intensities of tone. *J. genet. Psychol.*, 1937, **51**, 279–291.
10. HOVLAND, C. I. The generalization of conditioned responses: IV. The effects of varying amounts of reinforcement upon the degree of generalization of conditioned responses. *J. exp. Psychol.*, 1937, **21**, 261–276.
11. HULL, C. L. *Principles of behavior*. New York: D. Appleton-Century, 1943.
12. HULL, C. L. The problem of primary stimulus generalization. *Psychol. Rev.*, 1947, **54**, 120–134.
13. PAVLOV, I. P. *Conditioned reflexes*. (Trans. by G. V. Anrep) London: Oxford Univ. Press, 1927.
14. RAZRAN, G. Transposition of relational responses and generalization of conditioned responses. *Psychol. Rev.*, 1938, **45**, 532–538.
15. RAZRAN, G. Stimulus generalization of conditioned responses. *Psychol. Bull.*, 1949, **46**, 337–365.
16. SPENCE, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1937, **44**, 430–444.

(Received May 13, 1954)

Journal of Experimental Psychology
Vol. 49, No. 2, 1955

THE EFFECT OF LEARNING VERBAL LABELS FOR STIMULI ON THEIR LATER DISCRIMINATION [1]

JOHN S. ROBINSON

Cornell University[2]

Most of the experiments providing tests of Dollard and Miller's "acquired distinctiveness of cues" hypothesis (**3, 7**) have used motor discrimination criterion tasks (**2, 4, 5, 9**). The only study that provides data on the effects of "distinctiveness" training on a perceptual discrimination task is that of Arnoult (**1**), in which negative results were obtained. The data, however, suggest that more intensive preliminary training would result in significant facilitation of criterion-task discrimination. The present study is designed to test the distinctiveness hypothesis using a perceptual criterion task. The *Ss* were given intensive preliminary training, and a control group was also included to provide for evaluation of stimulus familiarization during preliminary training. In addition, a test was made of Dollard and Miller's "acquired equivalence of cues" hypothesis.

METHOD

Apparatus.—Ten fingerprints were selected as the experimental stimuli (**6**, Plates 1–4, 9, 11–15) and made into 2- × 2-in. slides. The *S* sat at one end of a table 18 in. from a partition that shielded the stimulus-presentation apparatus from his view. The fingerprints were projected from the rear on a 3½- × 2-in. translucent screen in the center of the partition. A tape recorder was used to record verbal responses during preliminary training. For the criterion task *S* made "same" or "different" judgments by pressing keys operating polygraph pens which recorded accuracy and latency of response.

Design of experiment and procedure.—University students and administrative personnel served as *Ss*. Three groups (*N*'s of 14, 15, and 16) were given different preliminary training tasks. Three *Ss* were dropped from the experiment for failure to reach the learning criterion, thus leaving 14 *Ss* in each of the three groups. These 42 *Ss* and an additional group of 14 *Ss* with no prior training were then given a criterion task.

Preliminary tasks.—The Distinctiveness Group (Group D) *Ss* learned distinctive verbal responses (gangsters' nicknames: Moll, Duke, Slim, Nick, Hymie, Butch, Rocky, Schnozz, Babe, and Squint) for the ten fingerprints by the paired-associates method. The *S* had access to a list of these names. A print was presented for 2 sec.; when it was removed from the screen, *E* pronounced its name. After the first trial, *S* had to give the name while the print was being shown. There was an 8-sec. interval between exposures and an intertrial interval of 45 sec. Trials were run to a criterion of one errorless trial. The *Equivalence Group* (Group E) *Ss* learned to call five of the prints "cops" and the other five "robbers" (arbitrary designations by *E*). The paired-associates method was used in the same manner described above. The *Ss* in Group E usually reached criterion (one errorless trial) in fewer trials than *Ss* in Group D. In order to equate *Ss* in the two groups with respect to number of trials, it was necessary in some cases to give an *S* in Group E several post-criterial trials. The *Sameness-Difference Group* (Group S-D) *Ss*' task consisted of ten stimulus presentations per trial, as in the above procedures, but a particular print might appear alone, or two or three times in a row. The *S* compared each print with the preceding one and said whether it was "same" or "different." Most *Ss* responded with 100% accuracy after the second or third trial so no learning criterion was necessary. The *Ss* in this group were matched with *Ss* in the other two groups by giving them the same number of trials. This

[1] This paper is part of a thesis presented to the faculty of the Graduate School of Cornell University for the degree of Doctor of Philosophy. The author expresses his gratitude to Drs. Eleanor J. Gibson and James J. Gibson for their continued aid during the conduct of the research and preparation of the thesis.

[2] Now at Yerkes Laboratories of Primate Biology.

task was included to control for any effect that prior opportunity to see and examine the prints might have on criterion-task performance, apart from any effects produced by learning responses for them.

Criterion task.—There were ten criterion-task trials. For each trial one of the ten fingerprints, used as a standard, was presented for 2 sec. followed by a series of ten comparison prints (each of .5-sec. duration) appearing at 8-sec. intervals. The *S* judged comparison fingerprints as same or different by pressing one of two keys in front of him. Five of the comparison prints were the same as the standard (the same slide was used) and the other five were selected from the remaining nine prints. "Sames" and "differents" were varied in a predetermined random sequence. Those *Ss* with preliminary training were told that the same prints would be used; *Ss* in Group NPT (no prior training) were instructed on fixating the screen, shown a fingerprint, and told that the experiment was "to determine their ability to discriminate fingerprints." All *Ss* were urged to make their judgments as rapidly and as accurately as possible.

Introspective reports were obtained from all *Ss* at the end of the criterion task.

Results

Preliminary training.—Group D required a mean of 18.40 trials to reach criterion with an *SD* of 7.14. Group E reached criterion in a mean of 12.31 trials (*SD* = 7.05). There was no criterion for the Group S-D task because performance reached 100% accuracy by Trial 2 or 3. Because all three groups were equated with respect to total number of trials given, total (criterial plus postcriterial trials in the case of Group E) trial means and *SD*'s were the same for the three groups.

The mean number of errors to criterion for Group D was 51.70 (*SD* = 30.63). For Group E the mean number of errors to criterion was 24.44 with an *SD* of 15.63. Errors on postcriterial trials increased the Group E values only slightly, most *Ss*' performances being error free during these trials. The *Ss* in Group S-D made a mean of 8.80 errors (*SD* = 6.18) during preliminary training. These *Ss* made nine judgments per trial and received a mean of approximately 18 trials (*SD* = 7.14). These error values, therefore, are based on a mean of 162 judgments, with some *Ss* making over 300 judgments.

Criterion task.—Table 1 shows the mean error and latency scores on the criterion task. Error score analysis of variance yielded an *F* ratio of 5.89 which is significant at the .01 level for 3 and 52 *df*. Only the differences between the three preliminary training groups and Group NPT were significant by *t* test. Latency score analysis of variance gave an *F* ratio of 1.34, which is not significant for 3 and 52 *df*.

TABLE 1

Means and *SD*'s of Error and Latency Scores on the Criterion Task
(*N* = 14 for each group)

Group	Errors*		Latency†	
	Mean	*SD*	Mean	*SD*
D	4.34	.79	33.04	11.48
E	4.30	1.14	35.25	12.22
S-D	3.88	1.17	31.32	7.21
NPT	5.15	.79	39.82	8.66

* A transformation ($\sqrt{x} + .5$) was used on the raw error scores to achieve homogeneity of variance.
† Multiply values in table by .02 to get reaction latencies in seconds.

Introspective reports.—These reports may be summarized as follows: (*a*) Reports on the preliminary training stage most frequently were concerned with *Ss*' attempts to isolate identifying characteristics in the fingerprints; (*b*) Some *Ss* reported using these identifying characteristics in making criterion task judgments; and (*c*) No *S* reported being aided or hindered on the criterion task by the verbal labels learned for the prints.

Discussion

The three preliminary tasks all had approximately the same facilitating effect on criterion-task performance. There was no enhancement of discriminability or equivalence as a consequence of learning responses for criterion-task stimuli. The results, therefore, do not support either of Miller and Dollard's hypotheses.

All three tasks involved (*a*) equal exposure to the stimuli, (*b*) attention to the stimuli, and (*c*) active search for identifying features in the fingerprints. These similarities among the tasks suggest an interpretation of the uniform transfer produced by the three tasks. All of the tasks provided the opportunity and necessity for developing *identifications* for the fingerprints (the finding of significant details enabling *S* to recognize a print and distinguish it from others). Errorless performance of the assigned tasks waited upon their development. The similarity in transfer shown by the three groups indicates that the learning of the arbitrary names for the prints by Groups D and E did not produce any further change in stimulus discriminability.

Such general factors as "warm-up" or "learning how to learn" may have produced the obtained transfer. Arnoult's results indicate, and detailed analysis of the criterion task on the present experiment (**8**) suggests, that these or other general factors do not contribute significantly to transfer. A control group is necessary, however, for more precise evaluation of the role played by such factors.

An additional problem needing clarification is the discrepancy between this study and that of Goss (**5**). Using a motor rather than a perceptual discrimination criterion task, and controlling for stimulus familiarization, Goss found that learned responses as such significantly facilitated discrimination.

Summary

Dollard and Miller's hypotheses of acquired distinctiveness and acquired equivalence of cues were tested. The *S*s were given either distinctiveness, equivalence, or stimulus familiarization training with a set of stimuli, and were then given a perceptual discrimination task using the same stimuli. A fourth group was given the discrimination task with no preliminary training. There were no significant differences in discrimination performance for the three preliminary training groups, and therefore the hypotheses are not supported. These three groups, however, made significantly fewer errors than the control group. Facilitation of discrimination on the criterion task is interpreted in terms of the development of identifications for the stimuli during preliminary training. General situational factors, however, may have produced the facilitation.

References

1. Arnoult, M. D. Transfer of predifferentiation training in simple and multiple shape discrimination. *J. exp. Psychol.*, 1953, **45**, 401–409.
2. Baker, K. E., & Wylie, R. C. Transfer of verbal training to a motor task. *J. exp. Psychol.*, 1950, **40**, 632–638.
3. Dollard, J., & Miller, N. E. *Personality and psychotherapy*. New York: McGraw-Hill, 1950.
4. Gagné, R. M., & Baker, K. E. Stimulus pre-differentiation as a factor in transfer of training. *J. exp. Psychol.*, 1950, **40**, 439–451.
5. Goss, A. E. Transfer as a function of type and amount of preliminary experience with task stimuli. *J. exp. Psychol.*, 1953, **46**, 419–428.
6. Henry, E. R. *Classification and uses of fingerprints*. (5th Ed.) London: Harrison, 1922.
7. Miller, N. E. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. abnorm. soc. Psychol.*, 1948, **43**, 155–178.
8. Robinson, J. S. An investigation of the effect of the learning of naming responses for stimuli on their later discrimination. Unpublished doctor's dissertation, Cornell Univ., 1953.
9. Rossman, I. L., & Goss, A. E. The acquired distinctiveness of cues: the role of discriminative verbal responses in facilitating the acquisition of discriminative motor responses. *J. exp. Psychol.*, 1951, **42**, 173–182.

(Received April 5, 1954)

Journal of Experimental Psychology
Vol. 49, No. 2, 1955

ADAPTATION TO TIME IN A SENSORIMOTOR SKILL[1]

R. CONRAD

Applied Psychology Research Unit, Cambridge, England

In his paper on the human factors governing biomechanical systems, Bentley (1) drew particular attention to "the temporal modes of activity." He pointed out that situations involving rapid change demanded the use of "extremely delicate psychological resources."

Although much knowledge of timing has been gained through classical studies of time perception, it has not proved easy to apply that knowledge to problems of biomechanics. To some extent this may be because in the majority of studies perceptual behavior was observed under somewhat abstract conditions rather than within the framework of continuously changing man-machine relationships.

In view of the elusive character of the primary variable, it is not surprising that the more realistic laboratory studies of the temporal modes of activity have on the whole been restricted to those relevant only to tracking. But, as Bentley indicates, the man in the system is liable to receive cues and signals from many different sources. In successfully dealing with them, timing will be as important as it is in dealing with the single series of cues in tracking.

In previously reported experiments (**2, 3**) the writer noted, but without discussion, that in dealing with a continuously changing display made up of four concurrent series of signals, the temporal accuracy of response appeared to be constant irrespective of the rate of display change.

It was pointed out (**3**) that not all signals were responded to, and missed signals were shown to occur close to a response more often than could be expected by chance. With the rates of change of the display used, the majority of signals were responded to. Since in matching his series of responses to the presented series of signals *S*'s temporal error was present even when the time which was available to formulate and execute response was relatively long, two questions presented themselves for further study: (*a*) Why under the most favorable temporal conditions was response in general apparently inaccurate? (*b*) Why was temporal accuracy apparently unaffected when signals were presented more rapidly? The present paper completes the analysis of the results of the originally reported experiment.

METHOD

Apparatus.—The apparatus and the experimental method have been described in the previous experiments (**2, 3**). Briefly, the display consisted of four dials each having six equidistant marks round the perimeter, and carrying a revolving pointer.

The *S* was provided with four Morse keys and was required to hit the appropriate key whenever a pointer coincided with a perimeter mark. This coincidence was regarded as a signal. The four pointers moved at slightly different speeds such that the relative time interval between successive signals on each dial was 1.0, 1.09, 1.40, 1.67. The effect of these differences was to present to *S* a series of signals, the temporal characteristics of which have been discussed by Cox and Smith (**4**). The distribution of the interval between successive signals considering the display as a whole approximated to the negative exponential. The occurrence of signals and *S*'s response to them were recorded on continuously moving paper.

[1] The work was carried out under the general direction of Sir Frederic Bartlett, C.B.E., F.R.S., and Dr. N. H. Mackworth, to both of whom the author is deeply indebted.

Procedure.—Six naval ratings acted as *S*s. They practiced for 15 min., morning and afternoon, for two days. Three experimental sessions then followed during which each *S* did 3, 2, and 2 tests, respectively, each of 5-min. duration. The following seven conditions of signal speed were given in random order: 40, 60, 80, 100, 120, 140, 160 signals/min. Each experimental session was preceded by a 3-min. "warming up."

Criteria of performance.—In the classical reaction time experiment, response is inevitably late. Normally it can only occur after the stimulus. Where "receptor anticipation" (5) is possible, response may occur either before or after the signal. Then if any single response is considered merely in relation to the signal which initiated it, no information is present to explain why the response was made early or late. The simplest explanation is to regard each response as one might regard the throw of a dart, and consider that its position will depend entirely on the degree of eye-hand coordination achieved.

Alternatively, one can regard the signal-response element not as discrete and isolated from other behavior, but as one element in a continuum of perception and response, following events, and followed by other events in definable temporal sequence. In this case the temporal position of a response in relation to the signal which initiated it might be expected to depend on the position of some previous event, and perhaps on some succeeding event as well. Furthermore, if evidence in support of this latter alternative can be adduced, then it would be logical to examine the extent of the validity of the hypothesis in relation to signal series of varying mean speeds.

Two recorded events were postulated as important determinants of the position of a response (R_2). The first was the previous response (R_1), considered in relation to the signal (S_2) initiating R_2. This was because R_1 would be the last event over which *S* had control before R_2 was made. It was considered that the longer the time interval R_1-S_2, the greater the probability that R_2 would be early.

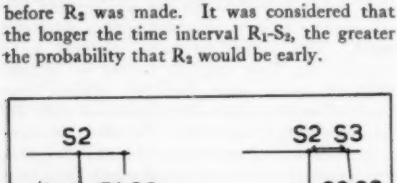

FIG. 1. The two intervals R_1-S_2 and S_2-S_3 exemplified.

TABLE 1

EFFECT OF SPEED CHANGES ON ACCURACY OF RESPONSE

	Speed in Signals/Min.						
	40	60	80	100	120	140	160
Accuracy (Sec.)	.26	.26	.25	.24	.25	.24	.24

The second event was the next signal (S_3) considered also in relation to the signal (S_2) to which the response (R_2) was made. This was because S_3 would be the event most worth considering if *S* in fact considered the future while dealing with S_2. It was postulated in this case that the longer the interval S_2-S_3, the greater the probability that R_2 would be late. These two critical intervals are illustrated in Fig. 1. Thus the definable temporal sequence would consist of a signal and response, the previous response, and the next signal.

Data were obtained from a 2-min. sample from each of the six *S*s. Performance at three signal speeds was examined: 40, 100, 160 signals/min. These were the slowest, middle, and fastest speeds of the whole range. The cases where two or more signals were responded to in the wrong order were excluded. In such cases in effect S_2 was dealt with before S_1. Interpreting these cases would be unnecessarily complicated, and would add little to what could be learned from less ambiguous data. This has led to a loss of about 10% of the material. It should be pointed out, however, that S_2 sometimes occurs before R_1 when R_1 is a late response. The interval R_1-S_2 therefore will sometimes appear as a negative value.

RESULTS AND DISCUSSION

The general results have been reported elsewhere (3), but those relevant to the present analysis are given in Table 1.

The effect of time elapsed since the previous response.—The interval R_1-S_2 was measured in every case in the sample, and classified according to whether R_2 was made early or late in relation to S_2. Where the position of R_2 was doubtful, that is, when response was very accurate, alternate cases were classified as early and late.

At all speeds there were more late

TABLE 2

INTERVAL R_1-S_2: PERCENTAGE NUMBER OF OCCASIONS FOR EACH SIZE OF INTERVAL THAT R_2 WAS LATE

Signals/ Min.	Interval Size (.25 sec.)						
	−1	1	2	3	4	5	5+
40	100.0	93.7	70.8	54.1	42.8	33.3	35.4
100	100.0	94.8	74.7	48.0	40.0	27.6	33.3
160	100.0	88.8	74.4	50.0	41.3	39.4	25.2

responses than early ones when the interval R_1-S_2 was short; the opposite was true when R_1-S_2 was long. This is shown in Table 2, which gives, for each size of the interval R_1-S_2, the percentage number of times that R_2 was late.

The effect of time until the next signal.—To test whether a response was independent of succeeding events, the procedure was repeated for the interval S_2-S_3. Table 3 shows the percentage number of cases that R_2 was early or late for each size of the interval S_2-S_3.

This time it will be seen that at all speeds late responses were associated with long intervals. But whereas at the slowest speed predominantly early response occurred with intervals up to about 1.25 sec., at the fastest speed predominantly early response did not occur beyond an interval of .25 sec., although the proportion of late responses remained dependent on the size of the interval.

It will be noted that whereas for a given interval size the effect of the preceding event is in general independent of signal speed, this is not true for the event following the response. For most sizes of interval, the faster the speed the higher the proportion of late responses. In Table 2 the proportions at the slowest and fastest speeds are not significantly different at the 5% level for any interval. In Table 3 in a comparison of the slowest and fastest speeds the difference between the proportions of late responses is significant for all intervals.

It has been shown that whether a response is made before or after the signal which initiated it depends partly on the time interval from the signal to the previous response, and partly on the time from the signal to the next signal responded to. Response tends to be late if the former interval is short or if the latter interval is long. At fast speeds, response will be likely both to be preceded and followed by short intervals. Similarly at slow speeds, there will tend to be long intervals before and after response, the former conducive to early, and the latter to late response.

Timing of R_2 in relation to previous and future time intervals taken together.—Certain inferences would seem to follow from the results of the analysis so far: (*a*) It should be possible to define optimum temporal conditions under which almost all responses will be made late, and others under which almost all responses will be made early. (*b*) It has been pointed out that conditions occur when both early *and* late response would simultaneously be

TABLE 3

INTERVAL S_2-S_3: PERCENTAGE NUMBER OF OCCASIONS FOR EACH SIZE OF INTERVAL THAT R_2 WAS LATE

Signals/ Min.	Interval Size (.25 sec.)					
	1	2	3	4	5	5+
40	20.0	27.3	29.6	42.9	42.1	52.7
100	27.4	50.0	46.3	61.4	58.8	65.9
160	40.8	54.3	51.3	59.6	63.2	65.2

expected. Since this is absurd, it should be possible to examine the conditions under which either opposing factor prevails. That is, if the two factors are such that they tend toward equal and opposite effects on the position of the response, under which conditions will the response in fact be early, and under which will it be late?

The analytical procedure was to examine every response (R_2) in terms of whether it was early or late, and in relation to the size of both time intervals R_1-S_2 and S_2-S_3. To avoid oversubdivision of the data, for the three speeds of 40, 100, and 160 signals/min., both intervals were classified into three groups as follows:

R_1-S_2	Short: up to .5 sec. (includes negative values) Medium: .5–1.0 sec. Long: Longer than 1.0 sec.
S_2-S_3	Short: 0–.5 sec. Medium: .5–1.0 sec. Long: Longer than 1.0 sec.

Table 4 gives the proportion of responses that were made *late* under each of the nine possible temporal conditions. Clearly, there are optimum conditions in which response will be predominantly late or predominantly early. If R_1-S_2 is short, and S_2-S_3 is long, some 90% of the responses will be made late. Similarly if R_1-S_2 is long and S_2-S_3 is short, from 60 to 80% of responses will be early. For a constant size of one interval, there is a fairly consistent relationship between the size of the other interval and the position of the response. But whereas in the case when one of the intervals is short the nature of the relationship is that expected, for some sizes of interval the relationship appears to be reversed. This shows primarily as a reversal of trend between the slowest and the fastest speeds when R_1-S_2 is long and the size of S_2-S_3 increases. The explanation for this is not obvious and it may be due to different effects of varying signal speed. In general, the consistency of the effects is good.

TABLE 4

EFFECT ON POSITION OF A RESPONSE (R_2) OF TEMPORAL STRUCTURE OF THE SEQUENCE AS DEFINED BY SIZE OF TIME INTERVALS R_1-S_2 AND S_2-S_3: PROPORTION OF LATE RESPONSES UNDER EACH CONDITION

R_1-S_2	S_2-S_3		
	Short	Medium	Long
40 Signals/Min.			
Short	.60	.67	.95
Medium	.26	.55	.54
Long	.21	.33	.40
100 Signals/Min.			
Short	.76	.93	.87
Medium	.36	.45	.49
Long	.23	.38	.38
160 Signals/Min.			
Short	.67	.82	.91
Medium	.48	.44	.43
Long	.40	.27	.22

If the events which influence the position of a response are equal distances away, will either prevail? If the proportions of late responses are compared when the intervals R_1-S_2 and S_2-S_3 are both short, both medium, or both long, it will be apparent that the common size of the two intervals affects the position of the response. At all speeds, if both intervals are short the majority of responses is late; if both intervals are long then the majority of responses is early. There are not enough cases of two short intervals at the

slowest speed for the significance of the difference between proportions to be reliably tested, but at the other two speeds there are significantly fewer late responses when both intervals are long than when both are short.

It has been suggested that for any one speed there is a balance of temporal forces around a response which does not depend only on the length of time between events. In other words, equal intervals of time on either side of a response change their resultant effect when the size of those intervals changes. One further suggestion relating to the interacting effects of speed and sequence characteristics on the position of response is implicit in the data of Table 4.

Consider in Table 4 the sequence consisting of long R_1-S_2 and short S_2-S_3. Such a sequence would be expected to result in predominantly early responses. Since the time intervals involved in this sequence are independent of signal speed, there is no reason to suppose, from the nature of the sequence itself, that the position of response will be affected by signal speed. There is, however, for the same sequence, a significantly greater proportion of late responses at fast speed than at slow speed. Other cases where speed significantly changes the quantitative effect of a sequence of standard characteristics are: medium R_1-S_2 and short S_2-S_3; long R_1-S_2 and long S_2-S_3. Some other changes will be noted but with the number of cases in the sample, they are not significant at the 5% level.

TABLE 5

EFFECT ON THE ACCURACY OF RESPONSE (R_2) OF TEMPORAL STRUCTURE OF THE SEQUENCE AS DEFINED BY SIZE OF TIME INTERVALS R_1-S_2 AND S_2-S_3: MEAN SIZE OF ERROR (SEC.)

R_1-S_2	S_2-S_3		
	Short	Medium	Long
40 Signals/Min.			
Short	.20	.25	.38
Medium	.21	.21	.21
Long	.30	.25	.25
	(Mean = .25)		
100 Signals/Min.			
Short	.18	.24	.36
Medium	.18	.21	.21
Long	.23	.23	.22
	(Mean = .24)		
160 Signals/Min.			
Short	.22	.27	.33
Medium	.20	.20	.23
Long	.26	.18	.26
	(Mean = .24)		

The size of error.—If the effect of certain changes in the temporal character of the sequence is to produce, for example, proportionately more late responses, then one would reasonably expect that the same changes would also result in a larger mean size of error between occurrence of signal and occurrence of response. With the same sequence classification as in Table 4—three speeds and three sizes of interval—the accuracy of each response was measured to the nearest .1 sec. Ignoring the plus or minus sign, the mean size of error in seconds for each sequence is given in Table 5.

Analysis of variance was carried out to examine (*a*) whether for any given speed there was an effect of sequence characteristic on accuracy, and (*b*) whether for any given sequence there was an effect of speed on accuracy. The result is shown in Table 6.

Table 6 shows that the sequence effect is highly significant while the difference between speeds for the same sequence is not. In other

TABLE 6

ANALYSIS OF VARIANCE: EFFECT ON ACCURACY OF SPEED AND SEQUENCE CHARACTER

Source	*df*	Square Mean	*F*	5%	1%
Speed	2	.00095	2.07	3.63	
Sequence	8	.00745	16.20		6.19
Remainder	16	.00046			
Total	26				

words, not only does the sequence of which a response forms a part tend to determine whether the response is made early or late, but it also influences to what extent it is made early or late. Perhaps the most curiously interesting example of this influence is the effect on accuracy when for a constant size of the interval S_2-S_3, the interval R_1-S_2 increases. When this occurs, in almost all cases (Table 4) when R_1-S_2 is short the majority of responses are late, but when R_1-S_2 is long most responses are early. The position of the response can be said, at some stage, to have passed through zero. In such cases one might expect the accuracy of response to have been better when R_1-S_2 was of medium size than when it was *either* long or short. Table 5 confirms this in a large number of cases.

Now, when the signal speed is varied, the temporal character of the signal series changes, and the frequency with which particular sequences occur will be affected. At slow speeds, sequences will more frequently involve long intervals, whereas at fast speeds short interval sequences will be common. Table 6 has shown that different sequences carry, as it were, different accuracy values for the response. The relationship between these different values and the frequency with which different sequences occur must determine the over-all mean accuracy for the complete series. The frequencies with which the sequences occur at three speeds are given in Table 7. They are expressed as percentages of the total for each speed. The necessary arithmetic required to arrive at the final mean error need not be shown here. The means are shown in Table 5 and it will be seen that the three values are very similar.

The main point which has been demonstrated is that in the skill under discussion, the mean accuracy of response in time is a function of the frequency with which particular sequences occur, since the accuracy of an individual response depends at least partly on the temporal structure of the sequence to which it belongs. Whether or not any particular sequence occurs is not wholly predetermined. The temporal relationships within a sequence are set up by *S*'s interacting with the elements in the

TABLE 7

FREQUENCY OF OCCURRENCE OF SEQUENCE EXPRESSED AS A PERCENTAGE OF THE TOTAL

R_1-S_2	S_2-S_3		
	Short	Medium	Long
40 Signals/Min.			
Short	1.33	1.60	9.60
Medium	5.06	2.93	14.93
Long	10.40	13.86	40.27
100 Signals/Min.			
Short	7.53	6.71	18.17
Medium	9.06	11.29	13.19
Long	13.75	11.13	9.17
160 Signals/Min.			
Short	5.30	15.22	15.77
Medium	15.90	14.54	9.78
Long	10.87	7.61	5.03

display which are predetermined. But once the structure of the sequence is established through response at one end and anticipated signal at the other, *S* is forced to adapt to time limits which are then inherent in the immediate situation.

This experiment provides a description of at least one of the temporal modes of activity of man in biomechanical systems; it underlines the presence of one of those delicate psychological resources which Bentley (1) mentioned. In particular, the methods used for measuring this kind of performance suggest an advance over the more traditional techniques of investigating the conditions for timing in skilled behavior. In laboratory studies of human factors in biomechanical systems there is the danger that error will be recorded merely as a gross score or total time. Frequently the conditions which give rise to error are not amenable to observation and are available only through subjective report. This paper demonstrates the possibilities in detailed analysis of sequential structure continuously recorded.

Summary

An attempt was made to explain the temporal accuracy with which a series of responses could be matched to a series of signals, by reference to the continuously changing temporal relationships between these events. Although approximately half of the responses were made too soon and half too late, this distribution was not fortuitous. The temporal position of a response in relation to the signal which initiated it was influenced both by the position of the previous response, and by the position of the signal for the next response.

Not only did these events determine whether a response would be made early or late, but to a large extent they influenced the size of this error. Thus, while the character of the temporal structure of events around a signal determined the relative accuracy of the response, the mean signal speed in relation to the mean response rate determined the frequency with which particular temporal structures occurred. The underlying factors that actually determined the character of the response were shown to be masked when the effects on response of the signal speed variable were presented as mean accuracy values.

References

1. Bentley, M. Forecast, timing and other primary factors in the government of certain biomechanical systems. *Amer. J. Psychol.*, 1952, **65**, 329–345.
2. Conrad, R. Speed and load stress in a sensori-motor skill. *Brit. J. industr. Med.*, 1951, **8**, 1–7.
3. Conrad, R. Missed signals in a sensori-motor skill. *J. exp. Psychol.*, 1954, **48**, 1–9.
4. Cox, D. R., & Smith, W. L. The superposition of several strictly periodic sequences of events. *Biometrika*, 1953, **40**, 1–11.
5. Poulton, E. C. Perceptual anticipation in tracking with two-pointer and one-pointer displays. *Brit. J. Psychol.*, 1952, **43**, 222–229.

(Received May 21, 1954)

Journal of Experimental Psychology
Vol. 49, No. 2, 1955

EFFECTS OF PROLONGED EXPOSURE TO LOW TEMPERATURE ON VISUAL–MOTOR PERFORMANCE [1]

WARREN H. TEICHNER AND JOHN L. KOBRICK

Quartermaster Research and Development Center

In an earlier study Teichner and Wehrkamp (3) found that groups of *S*s who practiced for only 20 min. per day on the pursuit rotor, each in a different ambient temperature, showed an impairment in performance compared to an optimal temperature group; that the magnitude of impairment depended on the magnitude of difference between the particular temperature and the optimum. Inasmuch as performance under the non-optimal conditions never reached the level of the optimal condition group during the 75 trials involved in the experiment, it was not possible to determine whether the effect of temperature was on the rate or on the limit of acquisition of the skill. Should it be on the rate alone, then it could be said that individuals tend to acclimatize with respect to visual-motor performance. If, however, temperature should affect the final limit of skill, then it would not be expected that individuals ever reach the level of performance attainable under optimal temperature conditions. One purpose of the present study was to determine which of these two hypotheses is more tenable.

Another question left unanswered by the study cited above is whether the effect of temperature is on the measure of performance alone, or whether it also affects the learning process which is developed with the practice trials. Should it be the latter, it would be expected that a shift in temperature during the acquisition process would result in a gradual change in performance, thus reflecting the effect of the previous temperature on the learning process. On the other hand, if it were a quick change, then temperature might be conceived of as strictly a performance variable. The present study was also designed with the aim of providing an answer to this question.

Method

Five soldiers selected at random from the study cited above were used as *S*s. There was a 5-month interval between experiments. These men clothed only in shorts, T shirts, and socks lived in a constant temperature chamber for the entire course of the experiment. An exact regimen was established and maintained, with one exception to be noted below. Details of schedule and biochemical, physiological, and psychophysiological measurements which were made will be reported elsewhere. For our purposes it should only be noted that the measures reported herein were always obtained at the same time of day and were not in conflict with or related to the other measures taken.

The experimental design consisted of three consecutive experimental periods. In chronological order they were: a baseline period lasting 16 days during which the room temperature was maintained at 75° F., a cold period lasting 12 days during which the temperature was held at 55° F., and a recovery period lasting 13 days during which the temperature was again 75° F.

Every day throughout the entire experiment, *S*s were given 15 practice trials on the pursuit rotor. The details of apparatus and procedure have been described elsewhere (3). These trials were 20 sec. long and had an intertrial interval of 10 sec.

Results

A plot of the pursuit-rotor data for the entire experiment in terms of mean

[1] Presented to the 1954 meeting of the Eastern Psychological Association.

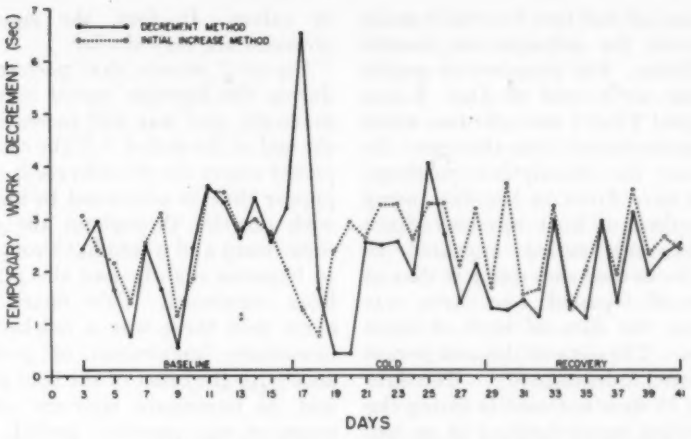

FIG. 1. Temporary work decrement as a function of practice under two different ambient temperatures.

time-on-target per trial indicated, among other things, that the initial practice portion of each day tended to be impaired as compared to the final level of the preceding day or the level attained on later trials of the same day. This phenomenon is known as temporary work decrement (TWD) and is characteristic of visual-motor tasks. In order to determine whether or not the effects of TWD were confounded with the experimental treatment, the magnitude of the decrement was calculated and plotted as shown in Fig. 1. Figure 1 shows TWD calculated by two methods: (*a*) as the difference in magnitude between the last trial of one day and the first trial of the next (decrement method), and (*b*) as the difference in magnitude between the first trial of a day and that trial of the same day at which the decrement appears to have been overcome (initial increase method). It can be seen that both methods show essentially the same things, a result which is consistent with previous findings (**2**). TWD although present showed no tendency to vary in magnitude with either practice or temperature except for the first day of the cold period. On this day the decrement method indicates a large increase in TWD, whereas the initial increase method shows a moderate decrease. This should be noted as an inversion in the order of magnitude of the two methods. Both curves appear somewhat cyclical in form, the cycles showing a lower frequency during the baseline period than during each of the other periods. Also, both curves show a marked change at the beginning of the cold period, each curve diverging in the opposite direction. Another interesting fact of Fig. 1 is that the decrement method tends to show a slightly but consistently smaller decrement than the initial increase method; this too is in agreement with previous results (**2**).

In order to obtain the most stable estimate of the treatment effect, the middle five trials of each day (Trials 6–10) were plotted as shown in Fig. 2. Omission of the first five daily trials eliminates the TWD and thus the artificial lowering of the mean scores.

Omission of the last five daily trials eliminates the influence of possible end effects. For purposes of graphing, the sixth trial of Day 1 was numbered Trial 1 and all other trials were renumbered from this one. In addition, for descriptive purposes, curves were fitted to the data using the method of least squares. Since the recovery period appeared to follow from the same trend as that of the baseline period, one curve was fitted to the data of both of these periods. The data of the cold period were fitted independently. The data of Day 19 were not used in fitting the cold period curve because of an unintentional rewarming of *Ss* on this day. It was found that an exponential curve provided the most reasonable descriptive function for both sets of values. In fact, the equations obtained are very similar.

Figure 2 reveals that performance during the baseline period increased gradually and was still increasing at the end of the period. If the recovery period scores are considered, it would appear that *Ss* continued to improve with practice throughout the entire experiment and would have continued to improve slightly had the practice been continued. This figure indicates that there was a marked and immediate impairment of performance with the onset of the cold period and an immediate recovery at the onset of the recovery period. Recovery of skill during the cold period was gradual and very similar in trend to that of the baseline period. The effects of a partial rewarming which

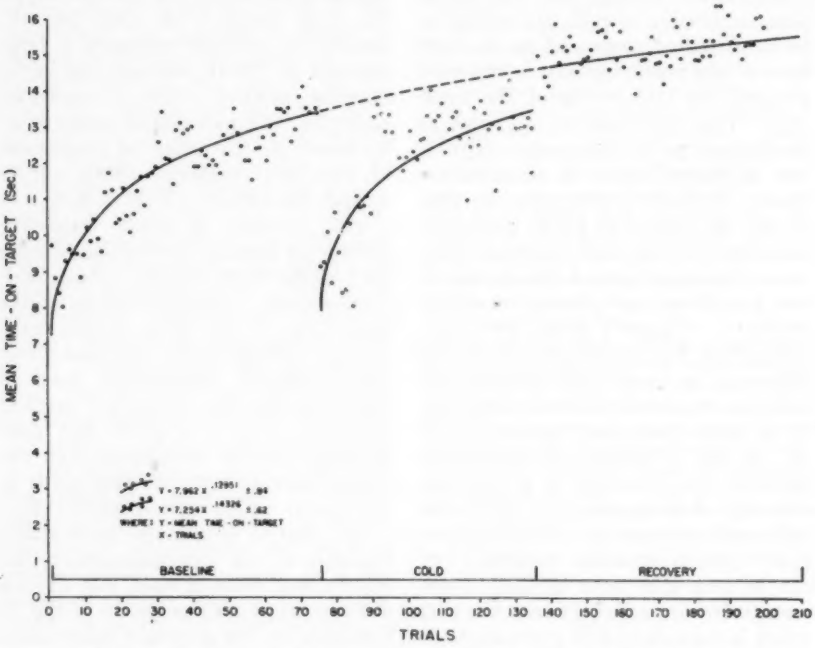

FIG. 2. Pursuit rotor performance as a function of practice under two different ambient temperatures.

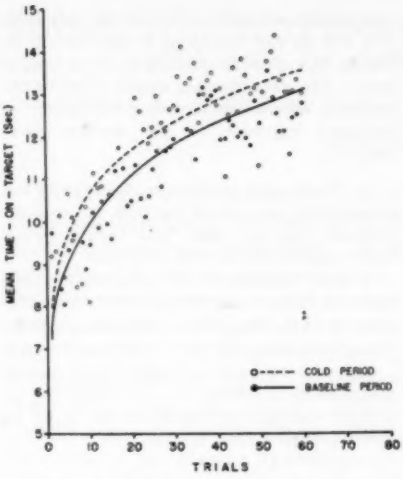

FIG. 3. Comparison of the effects of lowered ambient temperature on pursuit rotor performance with previous performance under optimal temperature conditions.

occurred on Day 19 are shown by the sudden rise in the plotted points of Trials 95 to 99. On these trials the scores almost approached final baseline performance.

The values used for the middle (dashed) portion of the over-all trend were obtained by calculation from the equation describing this trend. An *F* test was performed to determine the goodness of fit of these theoretical values to the data actually obtained during the cold period, including the data obtained on Day 19. An *F* of 13.75 was obtained from this test. With 58 and 236 *df* this is significant at the .001 level of confidence. It would appear reasonable, therefore, to conclude that the data of the cold period cannot be represented by the values which would be predicted for it by the over-all trend.

In order to determine more precisely the nature and amount of impairment which occurred in the cold, the actual cold period portion of Fig. 2 was plotted over the baseline period. This is shown in Fig. 3. This figure shows very clearly that performance during the cold period was reduced almost to the level of the beginning of the experiment; that recovery during the cold period closely paralleled baseline acquisition. In order to determine whether the cold period data could be represented by the baseline portion of the over-all trend, a test of the goodness of fit of the theoretical values of the baseline period to the cold period scores was performed. This yielded an *F* of 3.00 which, with 58 and 236 *df*, is significant at the .001 level of confidence. Thus, it may be stated that despite the closeness of the two sets of data, the cold period cannot be represented adequately by the baseline period of the over-all trend.

Discussion

The results support those of Teichner and Wehrkamp (3) in showing that there is an impairment in the cold, even in temperatures of 55° F. The results also show that the amount of impairment produced is very large, for not only was performance reduced to a level approximating the performance obtained at the beginning of the experiment, but the recovery from this level was slow and to the previous baseline level rather than to the level which might have been expected on the basis of practice alone. It is as if the effect of the lowered temperature had been to eliminate the beneficial effects of practice on the performance. Inasmuch as *S*s did not reach the expected level it cannot be said that they acclimatized, but only that they partly acclimatized. This being so, it can be concluded that low ambient temperature affects the final limit of visual-motor performance. Such a conclusion must be qualified by the possibility that had the cold period been extended, the *S*s might have recovered completely. This does not seem likely,

however, in view of the length of the cold period used and the trend of the scores during this period.

It is also clear both from the recovery displayed on Day 19 and from the quick shift in performance shown at the onset of the recovery period that temperature affects performance only and not the learning process. The rapid recovery to the expected practice level exhibited with the change to the recovery period would not have occurred if the conditions of the cold period had influenced the more permanent habit system.

The length of trial spacing and intertrial interval used are of an order such that they may be called relatively massed practice. The fact that the present study does not exhibit a decreasing trend in magnitude of temporary work decrement (TWD) with continued practice is consistent with recent data reported by Barch (2) and by Adams (1). Since the present study continued practice over an extended period of time and obtained the same results, there seems to be little doubt of the genuineness of the phenomenon. As to the influence of low temperature on TWD, there appears to be little or no relationship. At best there is a temperature effect only on the first day of exposure.

Summary

Five *Ss* lived in a constant temperature chamber for 41 days. For the first 16 days the temperature was held at 75°, for the next 12 at 55°, and for the remaining 13 days at 75° F. The *Ss* were given 15 trials daily on the pursuit rotor. The data obtained appear to allow the following conclusions regarding the effects of prolonged exposure to low ambient temperature:

1. Visual-motor performance is markedly and immediately impaired in the cold and recovers gradually, but to a lower limit than it attains under optimal temperature conditions.

2. The impairment of visual-motor performance in low temperature is the result of a lowering of the final limit of performance rather than a reduction of the rate or limit of learning.

3. Under conditions of massed practice temporary work decrement is not influenced by practice and affected slightly or not at all by lowered ambient temperature.

References

1. Adams, J. A. Warm-up decrement in performance on the pursuit-rotor. *Amer. J. Psychol.*, 1952, **65**, 404–414.

2. Barch, A. M. Warm-up in epicyclic pursuit performance under massed and distributed practice conditions as a function of the number of practice sessions. *USAF, Hum. Resour. Res. Cent., Res. Bull.*, 1952, No. 52-8.

3. Teichner, W. H., & Wehrkamp, R. F. Visual-motor performance as a function of short-duration ambient temperature. *J. exp. Psychol.*, 1954, **47**, 447–450.

(Received May 3, 1954)

Journal of Experimental Psychology
Vol. 49, No. 2, 1955

ROTARY PURSUIT PERFORMANCE AS RELATED TO SEX AND AGE OF PRE-ADULT SUBJECTS[1]

ROBERT B. AMMONS, STANLEY I. ALPRIN,[2] AND CAROL H. AMMONS

University of Louisville

Differences in performance of motor skills associated with age and sex have received considerable attention. Yarmolenko (**19**) reports increases in performance with increase in age for such skills as speed and exactness of walking, grasping and relaying objects, lying down and getting up, ball throwing, and standing jump. Similar results are reported by Jones (**11**) in his study of static dynamometric strength with children from 11 to 17 years of age, although the positive relationship does not hold for girls in the case of some tasks, such as thrust. In this case the curve rises somewhat and then falls again with an increase in age. Performance on a manual steadiness task, spool packing, and a serial discrimination task (**12**) also shows a positive relation with increase in age. Espenschade (**8, 9**) has studied performance of children from their tenth to their sixteenth year on the various parts of the Brace Test. She reports increases in performance with age for all subtests except target-throw for boys; for girls this relationship holds only for a few of the subtests, and for only a portion of the age range.

McCloy (**14**), studying push ups and pull ups, and Pyle (**17**), studying card sorting and digit-symbol substitution, obtained curves which have a clear upward trend in performance with increasing age. McCloy's age groups ranged from 12 to 16 yr., and Pyle's from 8 to 18 yr. Hicks (**10**), who worked with pre-school age children, found no age differences on several tasks. McGinnis (**15**) reported statistically significant differences in slot maze performance between 3-yr. and 5-yr.-old children, and between 4-yr. and 5-yr.-old children. For college students, there is no information about differences in rotary pursuit performance among different age groups, but there is some evidence that there are performance differences within age groups that are associated with sex (**6**).

It appears that the amount of gain in proficiency with age varies from task to task, and that in some cases, at least, performance level may decrease with increasing age. The present study was designed to investigate the relationships of sex and age to the continuous rotary pursuit performance of pre-adult *Ss* before and after a single rest period.

METHOD

Apparatus.—The basic rotary pursuit apparatus used here has been described in detail elsewhere (**4**), with the exception that only three of the four rotor units were utilized in the first experiment, but all four rotor units were used in the duplication of the experiment. The Koerth-type pursuit rotors consisted of black vinylite

[1] This research was supported in part by the United States Air Force under Contract No. AF 33(038)-10196, monitored by the Psychology Branch, Aero Medical Laboratory, Directorate of Research, Wright-Patterson Air Force Base, Ohio. Support was also received from the Research Committee, University of Louisville. The authors gratefully acknowledge the assistance of William Holley, Andrew Eckles, and Gene Farr in the collection and analysis of data, and critical reading by Drs. A. J. Latham and John Fink.

[2] Now with Bridgewater Township Public Schools.

turntables 10¾ in. in diameter, with ¾-in. brass targets set with their centers 3¼ in. from the turntable centers and flush with their surfaces. The *Ss* used light, hinged styluses (**5**) in following the target as it turned clockwise in a horizontal plane at 60 rpm. Two banks of 6-v. DC Standard Electric .001-min. timers were used alternately in the circuit to permit continuous recording of the duration of stylus-target contacts.

Subjects.—The total sample of 350 volunteer *Ss* was made up of 35 boys and 35 girls from each of five grade levels, 3, 6, 9, 11, and 12. Thirty *Ss* at each age level participated in the first experiment, and 40 *Ss* at each age level participated in the duplication of the experiment. Although the two experiments were run a year apart, *Ss* were from the same schools.[3]

The *Ss* were selected on the basis of approximately correct grade placement for age and lack of previous experience with the pursuit rotor. No *Ss* were taken from an experimental ninth grade typing class for boys, as the group was highly selective from the point of view of skill and would have volunteered as a whole if any had been allowed to serve as *Ss*. For the five grade levels, from 3 through 12, the mean ages and their *SD*'s for the boys were 9.11 (.69), 11.90 (.86), 15.03 (1.13), 16.83 (1.23), and 18.32 (.19), respectively; the means and *SD*'s of the girls' ages were 9.13 (.61), 11.66 (.21), 14.93 (.34), 16.71 (.68), and 17.68 (.97), respectively. The occupational status of the father or head of the family was recorded and tabulated according to the occupational classification of the U. S. Census (**18**). Breakdown of these totals for boys and girls within each of the five age groups indicated satisfactory similarity of parental occupation among the groups.

Procedure.—The *Ss* were run in groups of three in the first experiment and in groups of four in the duplication experiment. They were given the usual instructions: (*a*) to attempt to keep the stylus tip on the target as it rotated, (*b*) not to press down on the stylus, but to follow the target with a relaxed swinging movement of the arm, and (*c*) to relax the body to avoid becoming tired. Stylus grip and stance were demonstrated and corrected, and after a 3-sec. warning, the rotors were started. The *Ss* were encouraged periodically during practice.

Practice consisted of 3 min. of continuous rotary pursuit, followed by a 5-min. rest, and then 8 min. of continuous postrest practice. This particular work-rest-work cycle was used because it provides nearly optimum conditions for maximum performance (**2, 3**). Cumulative time on target was recorded at 20-sec. intervals. The duration of the intervals was timed manually with a stop watch. During the rest period *Ss* talked with *Es* and with each other.

The duplication experiment was similar to the original except that *Ss* were run in groups of four instead of three. The *Ss* were run by a different *E* a year after the completion of the original experiment.

Results

Since it was found that the results of the two experiments, when analyzed separately, were very similar, the data were combined. Although the main variables studied here are age and sex, the presentation of the results and their discussion are organized in terms of those phenomena which have figured in theoretical formulations (**2, 13**), in an attempt to discover clues to the acquisition process in rotary pursuit.

Performance before and after rest.—In Fig. 1 are plotted the mean time-on-target scores made by boys and girls at the several grade levels. In prerest practice the shapes of the curves for the two sexes and the five age levels are similar. However, with increasing age the curves rise more sharply, this rise being most marked for boys. The greatest slope appears in the prerest curve for the boys in Grade 12. Over-all greater increase in slope occurs in the performance curve of the boys.

The shape of the postrest performance curves also changes with increasing age. Children in Grade 3 show some gain over rest, this gain being somewhat greater for the boys. However, there is no identifiable warm-up decrement, and the decrement in subsequent performance is slight. Generally, the curves for both sexes are rather flat. By Grade 6 there is more marked gain over rest,

[3] The authors gratefully acknowledge the generous cooperation of Lucy Spurgeon and Ervin Detjen, principals of Cochran School, and S. V. Noe, principal of Girls High School, Louisville public schools.

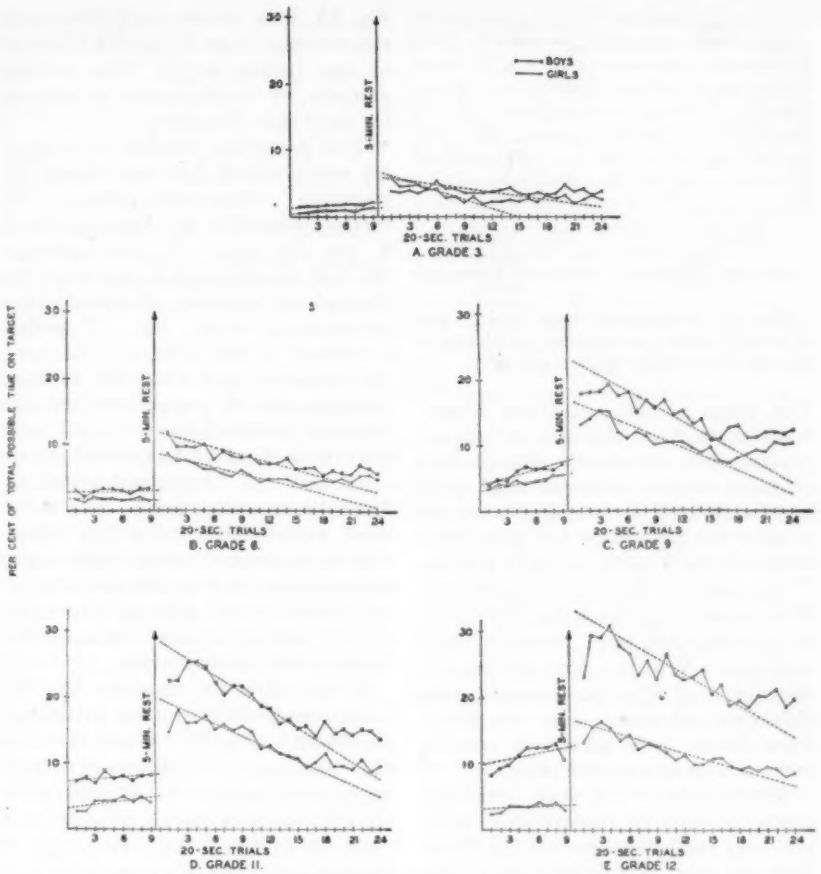

FIG. 1. Mean rotary pursuit performance before and after a single 5-min. rest as related to sex and grade level ($N = 35$ at each point).

followed by the usual decremental portion and gradual resumption of rise. The curves for boys and girls have quite similar shapes. The postrest curves for children in Grades 9, 11, and 12 show the typical postrest rotary pursuit phenomena (**2**) including warm-up decrement. However, inspection suggests that there are different amounts of reminiscence and warm-up decrement for boys and girls with increasing age. The curves shown for Grade 12 children appear to be very much like massed practice curves for college students (**3, 4, 6, 7**).

Level of performance is the third important aspect of acquisition. Inspection of Fig. 1 shows that with increasing age the initial level of prerest practice increases for both boys and girls, although more markedly for the former. Changes in final level of performance during both pre- and postrest practice are shown in Fig. 2.

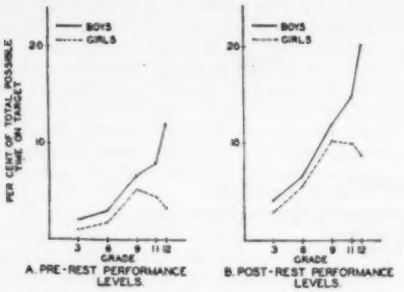

FIG. 2. Performance levels during last minute of prerest and postrest practice as a function of age (grade) for boys and girls.

The mean of the last three 20-sec. trials in each of the two periods of practice was obtained. Comparison of these curves indicates that with increasing age, performance level increases for boys; but for girls, performance level increases from Grades 3 through 9, and then decreases. This leads to an increasing difference in performance level between boys and girls. Whereas boys are always superior to girls in performance, this difference increases very markedly from Grade 9 to Grade 12 and is greatest during postrest practice.

Reminiscence.—Classical reminiscence, or gain in performance after rest over the level expected had there been no rest, was measured in the present study in the following manner. An estimate of the level which would have been attained had the group not been given a rest was obtained by forward extrapolation of a line drawn through the mean of the prerest performance curve. The slope was determined by the first and last points in this period. The difference between this point and the actual score made on the tenth or first postrest trial is the estimate of classical reminiscence. These estimates were made for boys and girls separately at each grade level. They are shown in Fig. 3A, from which it is evident that reminiscence is an increasing function of age (grade level). The over-all amount of reminiscence is greater for boys than for girls.

This particular method of measuring reminiscence does not correct for warm-up decrement; hence the method described by Ammons (**2, 3, 4**) was also used. He has suggested (**2**) that reminiscence is related to the dissipation over rest of a temporary decremental factor, D_{Wt}. This D_{Wt} is defined as the difference between the score estimated by the forward extrapolation of prerest performance level as described above to the next trial (here, Trial 10, had there been no rest) and the theoretical score on Trial 10 as estimated by the backward extrapolation of a line whose slope is determined by the highest and lowest points during postrest practice and which passes through the mean of the points included between the highest and lowest points.

Curves showing amounts of temporary work decrement so estimated are shown in Fig. 3C. These indicate that temporary work decrement (D_{Wt}) is an increasing function of grade level (or age) for both boys and girls from Grade 3 to Grade 11. In terms of percentage of time on target, D_{Wt} appears to be greater for boys than for girls; however this may be due to differences in performance level. It will be noted that generally the two

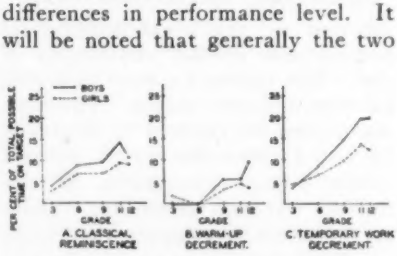

FIG. 3. Classical reminiscence, warm-up decrement, and temporary work decrement as functions of age.

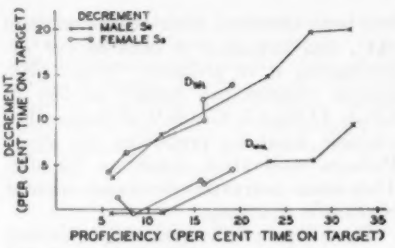

FIG. 4. Relationship between warm-up (D_{wu}) and temporary work decrement (D_{Wt}) and proficiency after rest, corrected for warm-up decrement.

methods of estimating reminiscence give rather similar results. Apparently when warm-up is not corrected for, the estimate of reminiscence (see Fig. 3A) is depressed.

Warm-up decrement.—A rough estimate of the amount of rise to a higher level after rest as related to age and sex can be made by comparing the curves in Fig. 1. However, a clearer picture is obtained in Fig. 3B, where the estimates were obtained in the way suggested by Ammons (**2**). He defined warm-up decrement (D_{wu}) as the difference between the actual score made on the first trial after rest (Trial 10) and the theoretical score arrived at by backward extrapolation of a line fitted to the highest and lowest points of postrest performance and drawn through the mean of the intermediate points. Estimates derived in this fashion and shown in Fig. 3B indicate that D_{wu} is also an increasing function of age, with no appreciable amount at Grades 3 and 6, but for boys an increasing amount with increase thereafter in age, and for girls an increasing amount at least through Grade 11.

DISCUSSION

Proficiency, temporary work decrement, and warm-up decrement all show systematic changes with age, and these changes are roughly similar for boys and girls. In Fig. 4, temporary work decrement (D_{Wt}) and warm-up decrement (D_{wu}) are plotted against proficiency (level of performance on the first postrest trial, corrected for D_{wu}), separately for boys and girls. There appears to be a basic relationship which is the same or very nearly the same for boys and girls. Temporary work decrement is directly proportional in amount to level of proficiency. Backward extrapolation of the curve suggests that the relationship may well hold back to zero proficiency. Warm-up decrement probably does not appear until a certain minimum level of proficiency is reached, and then develops more slowly than temporary work decrement. It should be noted that variations in proficiency here are not due to specific practice of rotary pursuit, since amount of practice was the same for all groups and *S*s were naive to the task.

The apparent relationships between D_{Wt} and D_{wu} and proficiency may be to some extent an artifact of the methods of measurement utilized. Units are probably "distorted" near the zero level of proficiency by some sort of threshold effect, with the result that D_{Wt}, D_{wu}, and proficiency vary concomitantly. It is also true that these three variables are not measured independently. An error in the estimation of proficiency after rest will correspondingly increase or decrease the estimates of D_{Wt} and D_{wu}.

Although measurement factors may increase the correlations among the measures, they probably do not account for more than a part of them. The results can also be interpreted speculatively in terms of transfer as follows. Basic proficiency is due to a combination of fundamental ability and transferred skill. Presumably the fundamental ability is about the same from age level to age level, so that changes in proficiency with age can be attributed to changes in amount of transfer. It is easy to see how variety of skills acquired and proficiency in them increase with age, this producing more transfer to "new" tasks. Since boys and girls are probably equal, or nearly so, in funda-

mental ability, the lesser proficiency of the girls and younger children is due to their having learned fewer skills and having learned them on the average less well than the boys and older children. With a greater number of habits from which transfer can take place, the boys and older children should experience greater negative as well as greater positive transfer. The "overcoming" of the interfering response tendencies leads to temporary work decrement, while their "spontaneous recovery" over rest leads to warm-up decrement. Thus we would expect to find greater warm-up and temporary work decrement associated with greater proficiency due to general transfer.

If the above speculation is at least partially correct, there should be corresponding correlated differences in warm-up and temporary work decrement associated with differences in proficiency within an age or sex group. That is, there should be more D_{wu} and D_{Wt} the greater the proficiency. Pomeroy (16) found that more proficient *S*s experienced a greater warm-up decrement after rest, but there seemed to be no appreciable difference in temporary work decrement. Adams and Reynolds (1) found no simple relationship among the three variables when they compared groups of trainee airmen differentiated according to initial level of ability. It is quite possible that fundamental ability is not related to D_{wu} and D_{Wt} in the same way that transferred skill is, and that differences in fundamental ability are complicating Adams and Reynolds' results.

The physical arrangement of the apparatus may have been responsible for some age differences in performance. Relative to the children's size and reach, the rotor turntable may have seemed larger for the younger children. However, the younger children did not seem to have difficulties associated with apparatus size, and, as a matter of fact, no child was so small that he or she could not easily reach the target when it was farthest away. In the present study the performance of the girls drops from Grade 9 to Grade 12. Since this drop has been observed with other activities (11), and appeared in both of our experiments, it is probably "real." Potential transfer can hardly be less in Grade 12 than in Grade 9, so some other variable must be producing the drop. Perhaps motivation sensitizes transfer. Then when motivation decreases, amount of transfer decreases.

The shape of the prerest acquisition curves in the present study may be of considerable importance. Not only does performance start at a higher level, but it also increases more rapidly in the case of the older children. Assuming that increase in proficiency with age is due to transfer, one can conclude that transfer influences not only the initial level of performance but also the rate of rise and final level of performance.

From a practical point of view it is worth pointing out that the relationship between age and proficiency is a continuous one, as are the relationships between proficiency and warm-up and temporary work decrement. It was also noted during the experimentation that younger *S*s cooperated as well on the whole as older *S*s. On these bases, there is little reason why various age groups and boys and girls cannot be used interchangeably in many experiments.

Summary

This experiment was conducted to determine the relationship for pre-adult male and female *S*s between age and various aspects of acquisition and performance of a simple motor skill. The *S*s were 350 children, 35 boys and 35 girls each in Grades 3, 6, 9, 11, and 12. Three minutes of continuous practice on the pursuit rotor were followed by a 5-min. rest, then a further 8 min. of continuous practice. It was found that: (*a*) there was a marked over-all increase in proficiency accompanying increase in age, with boys showing an increasing superiority over girls; (*b*) the proficiency of the girls declined from Grade 9 to Grade 12; (*c*) temporary work decrement (reactive inhibition) and warm-up decrement generally increased in amount with increasing age, and boys showed more than girls; and (*d*) temporary work decrement and warm-up decrement were roughly proportional in amount to level of proficiency, with temporary work decrement present at all levels, while

warm-up decrement did not appear until a certain minimum level of proficiency was reached. It was concluded that: (*a*) the results can best be accounted for in terms of a transfer theory of acquisition and decrement, which theory was outlined; (*b*) for most purposes *Ss* of both sexes and a wide age range can be used to study the phenomena of acquisition and decrement in rotary pursuit, even though there are marked quantitative differences in performance among the groups.

References

1. Adams, J. A., & Reynolds, B.[2] Rotary pursuit performance as a function of initial level of ability. *Amer. Psychologist*, 1952, **7**, 261. (Abstract)
2. Ammons, R. B. Acquisition of motor skill: I. Quantitative analysis and theoretical formulation. *Psychol. Rev.*, 1947, **54**, 263–281.
3. Ammons, R. B. Acquisition of motor skill: II. Rotary pursuit performance with continuous practice before and after a single rest. *J. exp. Psychol.*, 1947, **37**, 393–411.
4. Ammons, R. B. Acquisition of motor skill: III. Effects of initially distributed practice on rotary pursuit performance. *J. exp. Psychol.*, 1950, **40**, 777–787.
5. Ammons, R. B. A simple stylus for pursuit tasks. *Percept. mot. Skills Res. Exch.*, 1950, **2**, 22–24.
6. Buxton, C. E., & Grant, D. A. Retroaction and gains in motor learning: II. Sex differences, and a further analysis of gains. *J. exp. Psychol.*, 1939, **25**, 198–208.
7. Buxton, C. E., & Henry, C. E. Retroaction and gains in motor learning: I. Similarity of interpolated task as a factor in gains. *J. exp. Psychol.*, 1939, **25**, 1–17.
8. Espenschade, A. Motor performance in adolescence. *Monogr. Soc. Res. Child Develpm.*, 1940, **5**, No. 1.
9. Espenschade, A. Development of motor coordination in boys and girls. *Res. Quart. Amer. Ass. Hlth*, 1947, **18**, 30–44.
10. Hicks, J. A. The acquisition of motor skill in young children: an experimental study of the effects of practice in throwing at a moving target. *Univ. Ia. Stud. Child Welf.*, 1931, **4**, No. 5.
11. Jones, H. E. *Motor performance and growth: a developmental study of static dynamometric strength*. Berkeley: Univ. of California Press, 1949.
12. Jones, H. E., & Seashore, R. H. The development of fine motor and mechanical abilities. *Yearb. nat. Soc. Stud. Educ.*, 1944, **43**, 100–122.
13. Kimble, G. A. An experimental test of a two-factor theory of inhibition. *J. exp. Psychol.*, 1949, **39**, 15–23.
14. McCloy, C. H. The influence of chronological age on motor performance. *Res. Quart. Amer. Ass. Hlth*, 1935, **6**, 61–64.
15. McGinnis, E. The acquisition and interference of motor habits in young children. *Genet. Psychol. Monogr.*, 1929, **6**, 203–311.
16. Pomeroy, D. S. Retention of a motor act as a function of interpolated activity. Unpublished master's thesis, Univ. of Iowa, 1941.
17. Pyle, W. H. *Nature and development of learning capacity*. Baltimore: Warwick and York, 1925.
18. U. S. Bureau of the Census. *Sixteenth census of the United States, 1940, population*. Washington: U. S. Government Printing Office, 1940, *et. seq.*
19. Yarmolenko, A. The motor sphere of school-age children. *J. genet. Psychol.*, 1933, **42**, 298–318.

(Received May 24, 1954)

Journal of Experimental Psychology
Vol. 49, No. 2, 1955

AN ANALYSIS OF THE "VERTICAL–HORIZONTAL ILLUSION"

TEODOR M. KÜNNAPAS

University of Stockholm, Sweden

The vertical-horizontal illusion was one of the first illusions that was studied experimentally. In 1851 Fick (**4**) called attention to the discrepancy between vertical and horizontal judgments; he pointed out that "a bright square on a dark ground looks like a vertical oblong." In 1859[1] Wundt (**24**) published the first systematic study of that phenomenon. In this and a subsequent study (**25**) he tried to explain the tendency to overestimate a vertical line as compared with a horizontal line of equal length by eye movements. Helmholtz (**9**) observed the same illusion in connection with his lined squares.

Several earlier textbooks of psychology (e.g., **3, 17, 20**) and some recent ones (e.g., **1, 7, 13, 23**) mention the vertical-horizontal illusion, illustrating it sometimes by the **L** form; but usually the inverted **T** figure or a modification of this is employed. Sleight and Austin (**18**) investigated the overestimation of the vertical line in plane geometric figures.

Different attempts (**11, 12, 16, 19, 21, 22, 23, 24, 25**) have been made to explain the overestimation of the vertical. One of the most interesting hypotheses was put forward by Koffka (**10**). He pointed out that our phenomenal or behavioral space is anisotropic, and not Euclidean; and that it has different properties and stresses in different directions, even in its main directions such as horizontal and vertical. The overestimation of the vertical is a manifestation of this inequality of the horizontal and the vertical directions, and it appears in our perceptions.

It was Pan (**14**) who first emphasized that one of the vertical-horizontal figures—the inverted **T** figure—is an illustration of something more than the vertical-horizontal illusion. Finger and Spelt (**6**) found that the additional factor, included in the inverted **T**, is the bisection of the horizontal line by the vertical.

The present investigation involves study of different vertical-horizontal figures, not only the inverted **T**. It will be demonstrated that actually we have to do with two different illusions which produce the phenomenon that has, until now, been erroneously called the vertical-horizontal illusion.

One of these illusions is undoubtedly the overestimation of the vertical line in relation to the horizontal line of equal length. This classical illusion, however, cannot explain the whole effect. Figure 1 shows two horizontal lines (*b*) and (*c*), which are connected with a third vertical line (*a*), all three lines are of equal length. Nevertheless, one of the horizontal lines (*b*) appears shorter, and the other (*c*) longer than the vertical line (*a*). It is of particular interest to note that the horizontal line (*c*) appears longer than the vertical line of equal length (*a*). With regard to the overestimation of the vertical this phenomenon is inexplicable, almost paradoxical. The only possible explanation is to assume

[1] Although Wundt's monograph is dated 1862, parts of the Beiträge had appeared earlier in various issues of the *Zeitschrift für rationelle Medicin*.

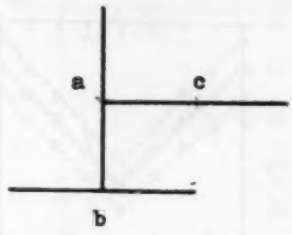

FIG. 1. The horizontal line (*c*) seems longer than the vertical line (*a*) of equal length.

the existence of a new illusion, which must be considerably stronger than the classical one, since it not only compensates for the overestimation of the vertical, but is even capable of effecting an obvious lengthening of the horizontal line.

The hypothesis underlying the present analysis is that the new illusion, which appears in the vertical-horizontal figures, is due to the overestimation of the dividing line. According to this hypothesis we overestimate the dividing line as compared with the divided line, irrespective of whether the direction is vertical or horizontal.

On the basis of this hypothesis, three testable predictions may be made. (*a*) If the dividing line is vertical and the divided horizontal, the two illusions are operating in the same direction, and therefore the size of the composite illusion is equal to the sum of the two illusions. (*b*) If, on the other hand, the dividing line is horizontal and the divided line vertical, the two illusions are acting in opposite directions, and the size of the composite effect is equal to their difference. If the two illusions are of equal strength, they will neutralize each other, and vertical and horizontal lines that are equal physically also appear to be equal. (*c*) Unlike the classical illusion, the amount of the overestimation of the dividing line is variable. The nearer the two parts of the divided line are to being equal, i.e., the more the position of the dividing line approximates to the midpoint of the divided line, the greater the illusion will be, and, on the contrary, the more unequal they are, the smaller the illusion will be.

METHOD

Stimulus materials.—The vertical-horizontal figures that were used were drawn in the center of white cardboard squares whose sides were 21 cm. Each square contained one figure. The length of the horizontal, divided line was 50 mm., while the length of the vertical, dividing line varied from 35 to 64 mm. The interval of variation was 1 mm. The lines were drawn with black India ink and their width was .3 mm.

The vertical dividing line intersected the horizontal divided at nine different points. These nine positions of the dividing line, one on each figure, were as shown in Fig. 2. As the length of the horizontal line was 50 mm., in the two last cases the dividing line touched the horizontal at its left and its right end, respectively. These two figures belong to the L form. The left direction will be indicated by a minus sign.

Procedure.—The exposure of the figures, one at a time, was orthogonal to *S*'s line of vision. The distance between *S*'s eyes and the figure was 100 cm. The surface of the square was illuminated by a lamp which was screened from *S*.

The constant method was employed in the experiments. After making preliminary trials, we selected, for every position, a limited number of squares which were presented to *S* in a pre-

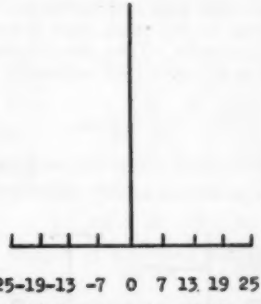

FIG. 2. Positions of the dividing line in relation to the divided line, in millimeters. Figure shows the midpoint position (inverted T figure).

arranged order that was unknown to him. The *S* had to compare the length of the exposed vertical and horizontal lines, and to report whether the dividing line is shorter, equal, or longer than the divided line. A maximum of 5 sec. was allowed for carrying out the judgments. The few cases where *S* could not make a definite report after 5 sec. were counted as "equal."

Experimental conditions.—Four conditions were arranged (see Fig. 3). In Cond. I the dividing line was vertical and directed upwards, and the divided line was horizontal. The way in which this experiment was conducted is described above. In Cond. II the squares containing the figures were rotated clockwise 90°, so that the dividing line was now horizontal, and directed to the right, whereas the divided line was in a vertical position. The *S* had to report whether the variable horizontal line was shorter, equal, or longer as compared with the vertical line. In Cond. III the figures were again rotated clockwise 90°, so that the dividing line assumed a vertical position, as in Cond. I, but was now directed downwards. The divided line was horizontal. The *S* had to report in the same way as in Cond. I. In Cond. IV the figures were again rotated clockwise 90°. Thus, the dividing line once again assumed a horizontal position, as in Cond. II, but was now directed to the left. The *S*'s task was the same as in Cond. II.

The purpose of Cond. I and III was to test prediction *a*. Conditions II and IV were designed to examine prediction *b*. The nine different positions in every experiment may also be used to test prediction *c*.

Each experiment was repeated four times on four different days. The sequence of the experiments was rotated so as to equate any possible learning and fatigue effects.

Subjects.—The *S*s were ten undergraduate students of psychology with normal visual acuity. In connection with the experiments that were conducted for four days, every *S* made about 1,300 judgments. Thus, the complete set of data is based upon 13,000 judgments.

Results

For each of the nine positions of the dividing line in every experiment, the point of the subjective equality (PSE) was taken as the significant measure of *S*s' responses. Also the upper limen (L_1) and the lower limen (L_s), as well as the interval of uncertainty (IU), were computed by Spearman's method of the arithmetic mean (**8**).[2]

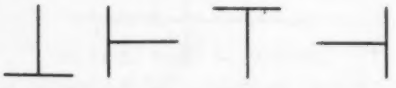

Fig. 3. Directions of the dividing line in the four conditions.

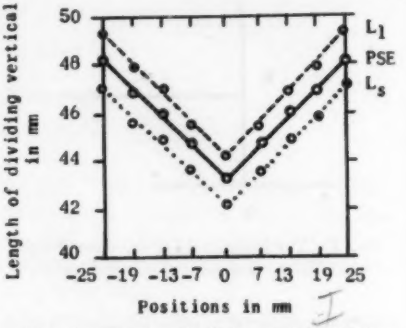

Fig. 4. The length of the vertical dividing line which appears equal to the horizontal divided line of 50 mm. The direction of the dividing line is upwards.

The dividing line upwards and downwards.—The results of Cond. I and III are very similar (see Fig. 4 and 5). The vertical dividing line is, in every position, shorter than the divided line, but the amount of the overestimation varies in different positions. The physical length of the vertical line, which appears equal to the horizontal divided line of 50 mm., is greatest at the end positions. The nearer we approach the midpoint position, the more the length of the vertical line diminishes. The diminution is rectilinear and symmetrical. The length of the vertical line is shortest in the midpoint position, i.e., the effect of the illusion has reached its maximum.

On analyzing these results, we dis-

[2] Complete tables of data by condition and variation of stimulus have been filed with the American Documentation Institute. Order Document No. 4425, $1.25 for microfilm or $1.25 for photocopies.

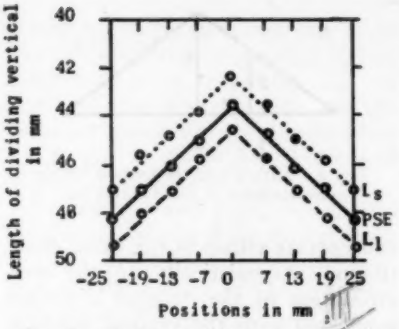

FIG. 5. The length of the vertical dividing line which appears equal to the horizontal divided line of 50 mm. The direction of the dividing line is downwards.

cover two factors operating in the same direction: (*a*) A *constant factor*. The average size of this factor is 1.71 mm. or 3.42%. It is the classical overestimation of the vertical line in comparison with the horizontal line of equal length; (*b*) A *variable factor*. It has its maximum value of 4.80 mm. or 9.6% at the midpoint position, and two minima at the end positions. At its maximum value the divided line is bisected; at the two minima, one part of the divided line is equal to the whole length of the line, whereas the other part equals zero, i.e., there actually is no division. Accordingly, the variable factor is here equal to zero. This factor is the overestimation of the dividing line which is due to the dichosection of one line by another. The effect of the total illusion is equal to the sum of the effects produced by the two factors.

The dividing line to the right and to the left.—The results of Cond. II and IV are also very similar (see Fig. 6 and 7), which leads to the conclusion that the direction of the dividing line is of no essential importance. Here the physical length of the dividing horizontal line varies in the same way as the dividing line in Cond. I and III, where it was vertical. The maximum of the illusion is also at the midpoint position, with two minima at the end positions. In this case, however, the dividing line is not always shorter than the divided line. On the contrary, at the two end positions it is actually longer than the divided line. Only on approaching the midpoint position it becomes first equal, and then shorter. This result

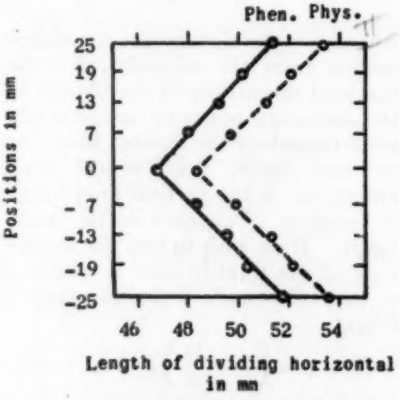

FIG. 6. The length of the horizontal dividing line which appears equal to the vertical divided line of the 50 mm. The direction of the dividing line is to the right.

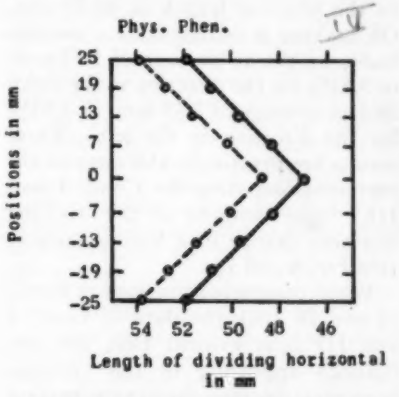

FIG. 7. The length of the horizontal dividing line which appears equal to the vertical divided line of 50 mm. The direction of the dividing line is to the left.

was to be expected since here the same two factors are operating antagonistically: the overestimation of the vertical lengthens the dividing line, which is here horizontal, and the dichosection shortens the same line. This means that the effect of the total illusion is here equal to the difference of the effects produced by the two factors.

When we compare the amount of overestimation of the vertical line in these conditions with the corresponding results of Cond. I and III, we can observe a considerable difference. In the latter case, the average overestimation of the vertical line was 1.71 mm., whereas here it is 3.64 mm. Yet it could be expected to be nearly the same in both cases.

However, it must not be forgotten, that here we had calculated with the physical length of the vertical line being 50 mm. The corresponding phenomenal length, owing to the overestimation of the vertical line, must be greater. When making a comparison, we must, however, take as our point of departure the apparent length of 50 mm., which corresponds to the physical length of 48.29 mm. On making a recomputation on this basis, we get an average of 1.77 mm. or 3.54% for the direction to the right, and an average of 1.83 mm. or 3.65% for the direction to the left. These results are practically the same as the corresponding data for Cond. I and III. Other lengths of the dividing line also change in a similar manner (see Fig. 6 and 7).

When comparing the data of Cond. II and IV with the data of Cond. I and III it is evident that the two illusions appearing in the vertical-horizontal figures—the overestimation of the vertical line, and the overestimation of the dividing line—are independent of each other: (*a*) They

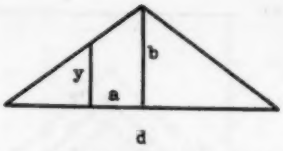

FIG. 8. Variables in the formula for the overestimation of the dividing line.

can operate either in the same direction or antagonistically; (*b*) the overestimation of the vertical is always connected with the vertical position, whereas the illusion due to the dichosection is independent of this direction. In the latter case, we overestimate the dividing line as compared with the divided line.

Magnitude of the illusion.—Since the overestimation of the dividing line increases or diminishes rectilinearly and symmetrically, it is not difficult to express its magnitude by a mathematical formula. From the two equal right-angled triangles (see Fig. 8) we have the equation $\frac{y}{b} = \frac{.5d - a}{.5d}$. Solving for y we obtain

$$y = \left(1 - \frac{2a}{d}\right) b, \qquad (1)$$

where a = the distance of the dichosection from the midpoint, b = the maximal magnitude of the illusion in the midpoint position, and d = the whole length of the divided line. As we saw above, the classical overestimation of the vertical line, which is constant c, appears in the same figure. If we wish to find the magnitude of the total illusion x, we have to add to formula 1 the constant c. Thus,

$$x = \left(1 - \frac{2a}{d}\right) b + c, \qquad (2)$$

where b and c are experimentally determined. We must bear in mind that the constant c may be positive or

negative. If the dividing line is in vertical position, c is positive, if it is horizontal, c is negative, i.e., it lengthens the divided line. At the midpoint position of the dividing line $a = 0$, and $x = b + c$. At the end position $a = .5d$, and $x = c$.

By using formula 2 we can compute the theoretical size of the total illusion in each of the nine positions and compare these with the empirical data obtained experimentally. Table 1 shows a comparison between the theoretical and the empirical data, where $c = 1.71$ mm. and $b = 4.80$ mm., which are the average results in the vertical position of the dividing line.

If we compute, by using the same b and c values, the theoretical data for the horizontal dividing line, and compare them with empirical data (see Table 2), we shall have another possibility of checking prediction c.

Thus, since all three predictions have proved to be valid, even in their quantitative formulation, they confirm our hypothesis.

Summary

The present study investigates the so-called vertical-horizontal illusion. Figures in addition to the L form and the inverted T were analyzed systematically. Four conditions were arranged, according to the four directions of the dividing line. In each condition nine different positions of the dividing line in relation to the divided line were used.

The principal findings are as follows:

1. In the vertical-horizontal figures two different illusions appear: (*a*) the classical overestimation of the vertical line as compared with a horizontal line of equal length, and (*b*) the overestimation of the dividing line. Together they produce the phenomenon that has, until now, been erroneously called the vertical-horizontal illusion.

2. The main differences of these illusions are: (*a*) The overestimation of the vertical line is always connected with the vertical position, whereas the illusion due to the dichosection is independent of this direction. Here we overestimate the dividing line as compared with the divided line. (*b*) Unlike the classical illusion, the latter is variable, it has a maximum at the midpoint position and two minima at the two end positions. (*c*) The amount of the overestimation of the dividing line can be considerably stronger than the overestimation of the vertical line, depending on the distance from the midpoint position.

3. Where the dividing line was vertical and the divided line horizontal, the two illusions operated in the same direction, and the amount of the total illusion was equal to their sum; where the dividing line was in a horizontal position and the divided line was vertical, the two illusions were acting in opposite directions, and the amount of the total effect was equal to the difference between their individual effects.

4. At the two end positions, where one part

TABLE 1

Comparison of Theoretical and Empirical Data of the Magnitude of the Total Illusion (in Millimeters) When the Dividing Line is Vertical

Positions	Theoretical	Empirical	
		Cond. I	Cond. III
−25	1.71	1.73	1.74
−19	2.86	3.04	2.98
−13	4.01	3.95	3.87
− 7	5.17	5.22	5.00
0	6.51	6.55	6.47
7	5.17	5.11	5.20
13	4.01	3.86	3.89
19	2.86	2.94	2.95
25	1.71	1.67	1.70

TABLE 2

Comparison of Theoretical and Empirical Data on the Magnitude of the Total Illusion (in Millimeters) When the Dividing Line is Horizontal

Positions	Theoretical	Empirical	
		Cond. II	Cond. IV
−25	−1.71	−1.85	−1.77
−19	− .56	− .57	− .76
−13	.59	.34	.42
− 7	1.75	1.70	1.79
0	3.09	3.12	3.06
7	1.75	1.83	1.65
13	.59	.56	.37
19	− .56	− .41	− .81
25	−1.71	−1.69	−1.88

of the divided line is equal to the whole length of the line (the L figure), the illusion due to the dichosection equals zero, and only the overestimation of the vertical line produces the illusory effect.

An attempt is made to express quantitatively the magnitude of the overestimation of the dividing line as well as the magnitude of the total illusion.

References

1. Bentley, J. E. *General psychology.* Philadelphia: J. B. Lippincott, 1947.
2. Boring, E. G. *Sensation and perception in the history of experimental psychology.* New York: Appleton-Century-Crofts, 1942.
3. Ebbinghaus, H. *Grundzüge der Psychologie.* II. Leipzig: Veit, 1913.
4. Fick, A. *De errore quodam optico asymmetria bulbi effecto.* Marburg: J. A. Kochii, 1851.
5. Fick, A. Erörterung eines physiologisch-optischen Phänomens. *Zeitschr. f. rationelle Medicin*, 1852, **2**, 83–94.
6. Finger, F. W., & Spelt, D. K. The illustration of the horizontal-vertical illusion. *J. exp. Psychol.*, 1947, **37**, 243–250.
7. Garrett, H. G. *Psychology.* New York: American Book, 1950.
8. Guilford, J. P. *Psychometric methods.* New York: McGraw-Hill, 1936.
9. Helmholtz, H. *Handbuch der physiologischen Optik.* III. 3. Aufl., Hamburg: L. Voss, 1910. (1. Aufl., 1866.)
10. Koffka, K. *Principles of gestalt psychology.* London: Routledge & Kegan Paul, 1950.
11. Lipps, T. Aesthetische Faktoren der Raumanschauung, in *Beiträge zur Psychol. und Phys. der Sinnesorgane*, published as *Festgruss* to Helmholtz, 217–307. Hamburg & Leipzig, Leopold Voss, 1891.
12. Lipps, T. Raumaesthetik und geometrisch-optische Täuschungen. *Schriften der Gesellschaft. f psychol. Forsch.*, No. 9 & 10. Leipzig: J. A. Barth, 1897,
13. Munn, N. L. *Psychology.* (2nd Ed.) Boston: Houghton Mifflin, 1951.
14. Pan, S. [The Vertical-Horizontal Illusion.] *N. C. J. Psychol. nat. cent. Univ.*, 1934, **1**, 125–128.
15. Pollock, W. T., & Chapanis, A. The apparent length of line as a function of its inclination. *Quart. J. exp. Psychol.*, 1952, **4**, 170–178.
16. Ritter, S. M. The vertical-horizontal illusion. *Psychol. Monogr.*, 1917, **23**, No. 4 (Whole No. 101).
17. Sanford, E. C. *A course in experimental psychology.* Boston: Heath & Co., 1898.
18. Sleight, R. B., & Austin, T. R. The horizontal-vertical illusion in plane geometric figures. *J. Psychol.*, 1952, **33**, 279–289.
19. Thiéry, A. Ueber geometrisch-optische Täuschungen. *Phil. Stud.*, 1896, **12**, 67–125.
20. Titchener, E. B. *Experimental psychology.* I. New York: Macmillan, 1901.
21. Valentine, C. W. Psychological theories of the horizontal-vertical illusion. *Brit. J. Psychol.*, 1912–1913, **5**, 8–35.
22. Valentine, C. W. The effect of astigmatism on the horizontal-vertical illusion, and a suggested theory on the illusion. *Brit. J. Psychol.*, 1912–1913, **5**, 308–330.
23. Woodworth, R. S. *Experimental psychology.* New York: Henry Holt, 1947.
24. Wundt, W. *Beiträge zur Theorie der Sinneswahrnehmung.* Leipzig & Heidelburg: C. F. Winter, 1862.
25. Wundt, W. Die geometrisch-optischen Täuschungen. *Abh. sächs. Ges. Wiss., math.-phys. Cl.*, 1898, **24**, 53–178.

(Received April 29, 1954)

Journal of Experimental Psychology
Vol. 49, No. 2, 1955

THE VARIATION IN WIDTH AND POSITION OF MACH BANDS AS A FUNCTION OF LUMINANCE [1]

CELESTE McCOLLOUGH

Olivet College

If two uniformly illuminated rectangles of different luminances are viewed one above the other, and in such a way as to have one side in common, an observer reports a contour at the common border. If, now, instead of being uniformly illuminated, one or both rectangles display a changing luminance from top to bottom, with a sharp but continuous change in luminance at the common border, *O* reports a band of finite width at this border, the band being either darker or brighter than the fields immediately contiguous to it. This sharp change may be described as a change in the rate of change of luminance. A negative change in the rate of change is correlated with the report of a bright band, a positive change with the report of a dark band.

Observation of this effect has a long history, but little in the way of systematic analysis has resulted. Mach must be credited with having first reported the observation (**5, 6, 7, 8, 9**) and with having provided the most complete analysis of its characteristics. A number of subsequent observers, including McDougall (**4**) and Thouless (**11**), have described the phenomenon independently.

[1] This report was prepared under Project No. NR 142-404, Contract N6onr-271, Task Order IX, between Columbia University and the Office of Naval Research, U. S. Navy. Reproduction in whole or in part is permitted for any purpose of the U. S. Government. The writer is indebted to Professor C. H. Graham for his guidance of the study here reported.

Since Mach's first report, it has not been necessary to question the dependence of these bands upon a sudden change in the rate of change of luminance, or to demonstrate their subjectivity. Photographs of rectangles varying appropriately in luminance show the band phenomenon. The absence of maxima or minima of density in the photographic plates may easily be verified by examining them with a densitometer. It is also clear that the phenomenon is ultimately dependent upon interaction, since progressive restriction of the surround above and below the band weakens it and eventually results in its disappearance (**9**, p. 647).

To account for the band phenomenon in terms of interaction effects, it is essential to have precise quantitative information about its dependence upon luminance characteristics of the stimulus field. In describing these characteristics, it is helpful to think of the stimulus field as divisible into unit areas, or into rows of such unit areas, whose luminance remains constant in one dimension (along the rows) and varies along the other dimension. The course of this variation will be referred to as the luminance gradient, and its characteristics may be stated mathematically or described graphically.

A type of luminance gradient giving rise to distinct bright bands was employed in the present investigation. The graph of this gradient (Fig. 1) shows a flat limb of constant maximum luminance, intersected by a

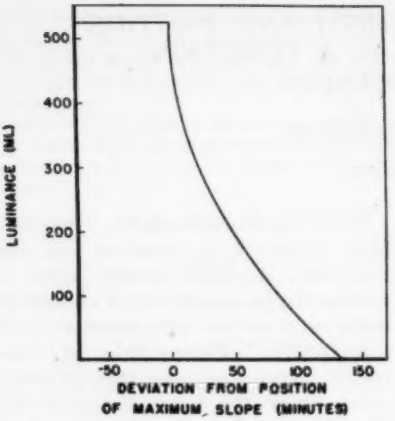

FIG. 1. Graph describing the form of the luminance gradient employed. The abscissa corresponds to the vertical dimension of the stimulus square; the left end of the abscissa represents the bottom edge of the square. The bright band appears in the vicinity of the point of maximum slope of the gradient, which is represented here at 0 on the abscissa.

concave limb describing a negatively accelerated decrease in luminance. The bright band is seen in the vicinity of the intersection, where change in luminance is most rapid.

Experimental operations on this gradient were of two kinds, which will be described in detail in the next section. One operation added a constant amount of light to all points in the stimulus field. Such an operation does not change the slope characteristics of the luminance gradient, but merely shifts the position of the gradient on the luminance axis. Luminance decreases at the same rate as previously, but it begins decreasing from a higher maximum luminance value. The other experimental operation produced a proportional increase or decrease in the luminance of all points in the stimulus field, changing the slope characteristics of the gradient as well as the maximum luminance value.

In discussing the effect of these experimental operations upon band width, it is convenient to treat the luminance gradient in terms of its logarithmic transformation. The log luminance gradient describes relative luminance change. Addition of a constant amount of light to all points affects the slope characteristics of the log luminance gradient, since it alters the relative luminance differences. A proportional increase or decrease in the luminance of all points in the stimulus field does not affect its slope characteristics, but merely shifts the position of the gradient with respect to the log luminance axis.

Two log luminance gradients with different degrees of curvature of the concave limb[2] were chosen for study. Each was observed at five positions on the log luminance axis. Measurements of width and position of the apparent bright band were made at each position and will be reported in relation to these variations in the log luminance gradient. The results will also be discussed in terms of the effect of varying position and slope of the luminance gradient itself.

METHOD

The measurement of width of band requires a means of producing suitable log luminance gradients, a method for shifting the position of the gradient on the log luminance axis, and a manner of indicating the position of the contours of the band. The nature of the situation in which these conditions were met will be indicated first; the apparatus used will then be described in detail.

The *O*'s eye receives light which originates at a uniform source and upon which a luminance gradient has been imposed by an intervening sector disc. Luminance is constant over the

[2] In the present report, the difference between the gradients will be specified only by graph, and the descriptive terms "steeper" and "shallower" will be used to refer to the gradients showing greater and lesser concavity, respectively.

FIG. 2. Stimulus square, as seen by *O*. Indicator line is not shown in the photograph.

lower half of the stimulus field; near the center, luminance decreases sharply, and continues to decrease at a decreasing rate between the center and top of the patch. The *O* sees a square field of steady light, surrounded by darkness, with its lower half brighter than the upper half and separated from it by a band brighter than either the upper or the lower half. He may indicate the upper and lower contours of the band by aligning a small black horizontal line segment with either of the contours. Figure 2 is a photograph of the stimulus square.

Figure 3 shows the apparatus schematically. The light source at *L* is a GE No. 213 photoenlarger bulb with an evenly coated white surface. It is operated on DC current at constant voltage to give a luminance of 7250 ml. Directly in front of *L*, at 18.25 in. from the eye, the sector disc (*D*) is mounted on the shaft of a ¼-hp. universal motor. The field stop (*FS*) is a 1-in. square at 15 in. from the eye, limiting the stimulus patch to 3.82° of visual angle. A glass plate (*P*) in contact with the field stop may be moved vertically by a micrometer screw; it carries the indicator line, 18′ by 1.6′, in a path in the approximate center of the stimulus square. Neutral density filters in the filter box (*FB*) control luminance of the field as a whole.

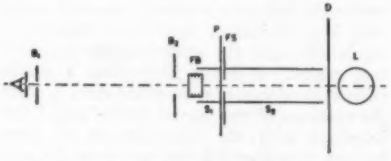

FIG. 3. Schematic diagram of apparatus.

Seated in a darkroom, *O* looks at the field through a 3-mm. artificial pupil in the left eyepiece of a binocular viewer. The right eyepiece is occluded, and the viewer is mounted on a wooden tube, 12 × 4 × 6 in., fastened to the wall of the darkroom. Masks (B_1, B_2) inside the blackened tube and shields (S_1, S_2) outside the darkroom insure zero luminance when the opaque sector disc passes before the light source.

The rotating disc is mounted in a position such that a section of the disc passing over the center of rotation intercepts the light from *L*. The shape of the sector cut from the disc determines the characteristics of the luminance gradient according to the Talbot-Plateau law. By that law, when luminance of the dark flash is zero, luminance at any distance *r* from the center of rotation is given by

$$B = B_L\left(\frac{\theta}{2\pi}\right),$$

where B_L is the bright flash luminance and θ is the size of the open sector in radians at the distance *r*.

Figure 4 is a diagram of the disc employed. The open sector is bounded by two radii, *N′P* and *QN*; by a straight line, *PQ′*, perpendicular to *OP* at *P*; and by arcs *QQ′* and *NN′*. A dotted square above *O*, the center of rotation, indicates the position of the stimulus field with respect to the disc. For convenience

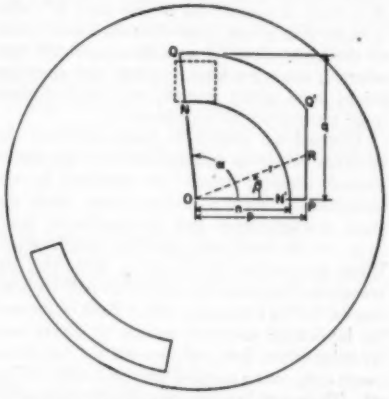

FIG. 4. Scale diagram of the gradient disc. Solid lines at the upper right indicate the critical open sector; lines at the lower left indicate a balancing sector. The dotted square shows the position of the stimulus field with respect to the disc.

of exposition, points in the stimulus field will here be described by the radial distance from O at which their light rays pass the sector disc.

When the disc rotates, luminance is the same for all rays passing the disc at a distance from O which is greater than n and less than p; this luminance is given by $B_L \frac{\alpha}{2\pi}$. For distances from O greater than p, the angular size of the open sector is less than α, by an amount β which varies as a function of the distance from O. Luminance of rays passing the disc at a distance r from the center, when $q > r > p$, is given by $B_L \frac{\alpha - \beta_r}{2\pi}$, where $\beta_r = \arccos p/r$. Differentiation of this function reveals that the rate of change of luminance is infinite when $r = p$ and decreases as r increases. This is also true of the first derivative of the logarithm of luminance.

By coupling a second disc with the first, covering all or part of the open sector between OQ and OQ', the size of α may be reduced. When this is done, a constant amount of light is subtracted from all points in the stimulus field. As α is reduced, variation in β as a function of r has a greater relative effect on luminance, and the log luminance of rays passing the disc at distances from the center greater than p decreases more rapidly with increasing distance from the center. The concave limb of the log luminance gradient is thus made steeper in the vicinity of the intersection of the two limbs. Log luminance gradients produced by a minimal and a maximal value of α were chosen for study. The value of α could not be less than 34°, with a sector disc of the dimensions used and a field of the given size. A maximum α of 96° was selected, since for greater values the apparent bright band which resulted was too indistinct to be measured with confidence.

Proportional change in luminance over the stimulus field was accomplished by the use of neutral density filters. As described in the introduction, such a change alters both the slope characteristics and the maximum luminance of the luminance gradient being studied. When the gradient is treated in terms of a log transformation, such changes affect only its position on the log luminance axis. Each of the two log luminance gradients spoken of above was investigated at five positions on the log luminance axis, over a range of approximately 3.5 log ml. These positions are conveniently designated by the logarithm of maximum luminance. The five values of log maximum luminance used were not identical for the two gradients.

Both gradient disc and cover disc were prepared from 12-in. diameter coated aluminum (phonograph) recording blanks. Dimensions of the gradient disc sector were, in inches, $n = 3$, $p = 3.5$, and $q = 4.5$. This sector was balanced by a sector of equal moment on the opposite side of the disc, placed beyond the visible area between arcs of radius 4.5 and 5.5 in. The cover disc had two opposite sectors of 100° each, bounded by arcs of radius 3 and 4.5 in.

In an experimental session, observations were made on a single gradient at one luminance level. The O first dark-adapted for 5 min. (10 min. at the two lowest luminance levels), then light-adapted for 1 min. to the luminance of the test field. He then adjusted the indicator alternately for the upper and lower contours of the band. Adjustments for position of the upper contour were always made by moving the indicator downward from points well above the band, until the lower edge of the indicator was aligned with the upper contour of the bright band. Adjustments for position of the lower contour were made upward from points well below the band, aligning the upper edge of the indicator with the lower contour of the band.[3] The O made 10 to 15 pairs of settings, proceeding at his own rate and closing his eye whenever necessary to rest it. Between adjustments he maintained approximate fixation at the center of the band. Ten such sessions, of 30 to 45 min. each, were required to obtain a set of data.

It was found in preliminary sessions that the variability of adjustments continued to decrease with practice during as many as ten sessions for some Os. With sufficient practice, adjustments could be made such that the SD of the width determinations in an experimental session was well under 12% of the mean of these determinations. When data obtained in any session showed a greater variability, they were not ac-

[3] It is of interest to note that the adjustment was necessarily an alignment, not a setting of the indicator edge to coincide with the contour of the band. The presence of the black indicator line in the vicinity of either border of the bright band led to disappearance of that portion of the band. The O could still indicate the position of the contour by aligning the edge of the indicator with those intact parts of the contour to the right and left of the disturbed region. The most probable explanation for this interference effect is that the indicator, a sharply bounded area of very low luminance, must have imposed irregularities upon the gradient being studied in that region. It seems likely that a similar effect may operate in some of the cases in which the presence of one border in the visual field interferes with the establishment of other borders. Cases of this sort have been discussed by Fry and Bartley (2).

cepted as sufficiently reliable, and the session was repeated.

Four *O*s, all emmetropic, gave data. One was a graduate student and three were undergraduates at Columbia or Barnard College.

At the end of each experimental session, *E* determined the objective position in the stimulus field of the point of maximum slope of the luminance gradient. This involved a series of adjustments for position of the point *P* on the gradient disc with the disc stationary. The disc was turned so that *PQ'* was horizontal and *N'P* coincided with the path of the left edge of the indicator. Then the lower edge of the indicator was aligned with *PQ'* in a series of ten settings; a bracketing procedure was used. The mean of these settings was taken as the objective position of *P*. The *SD* of these determinations was found empirically to be .15'. During a series of ten experimental sessions, the objective position of *P* so determined was found to vary as much as 3.0'.

Results

The data of this experiment are settings of the apparent positions of the upper and lower contours of the bright band. The width of the bright band may be derived from these data, as a measure of the band phenomenon. Width is the difference between a position setting of the upper contour and the next following position setting of the lower contour, minus the width of the indicator.

Changes in width of band have been investigated as a function of the logarithm of luminance at the point of most rapid change of luminance (for these gradients, log maximum luminance), for two gradients differing in rate of decrease of log luminance. Results are presented in Tables 1 and 2 and Fig. 5. For each log maximum luminance value, the tables give width of band (in minutes) for each of four *O*s for the steeper gradient, and for each of three *O*s for the shallower gradient. Each entry for an *O* is the average of several determinations; the number of determinations and *SD*'s are also shown. Figure 5 presents a comparison of widths obtained for the two gradients. Separate figures are presented for each of the three *O*s who gave data for both gradients. Curves have been drawn through the data by inspection.

The positions at which the contours of the band are seen are themselves of interest. These positions have been treated as deviations from the point of most rapid change of luminance (described here as the point of maximum slope), as determined at the end of each experimental session. Both the upper contour and the maximum slope positions were determined with reference to the lower edge of the indicator. Therefore, the deviation of the apparent upper contour from the point of maximum slope is the difference between the position setting for the upper contour and the

TABLE 1

Width (in Minutes) as a Function of the Logarithm of Maximum Luminance (Millilamberts) for the Steeper Gradient

log B_{max}	Subject EH			Subject CC			Subject BC			Subject AW		
	N	Mean	*SD*	*N*	Mean	*SD*	*N*	Mean	*SD*	*N*	Mean	*SD*
−0.80	12	12.2	1.2	14	11.7	0.88	12	13.1	0.69			
−0.56										15	10.8	0.82
−0.30	11	8.28	0.81	14	9.63	0.30				15	8.64	0.92
−0.13							10	9.99	0.56			
0.74	12	7.38	0.54	14	7.02	0.60	12	8.55	0.59	15	7.02	0.73
1.72	10	6.57	0.54	14	5.58	0.46	12	7.11	0.61	15	5.76	0.52
2.72	11	5.67	0.45	14	5.22	0.31	12	6.66	0.55	15	5.49	0.72

TABLE 2

WIDTH (IN MINUTES) AS A FUNCTION OF THE LOGARITHM OF MAXIMUM LUMINANCE (MILLILAMBERTS) FOR THE SHALLOWER GRADIENT

log B_{max}	Subject CC			Subject BC			Subject AW		
	N	Mean	*SD*	*N*	Mean	*SD*	*N*	Mean	*SD*
−0.56	12	11.7	1.2						
−0.33				10	15.4	1.1			
−0.11							15	11.2	1.2
0.10	12	9.18	0.58				15	10.1	0.91
0.32				10	12.8	1.2			
1.20	12	7.65	0.52	10	10.3	0.68	15	8.37	0.90
2.19	12	6.21	0.59	10	8.91	0.89	15	8.10	0.67
3.17	12	7.11	0.71	12	7.56	0.77	15	8.01	0.69

position of the point of maximum slope for that session. This difference indicates the extent of the bright band into the region of decreasing luminance, above the point of maximum slope. Since the apparent lower contour was indicated by the upper edge of the indicator, its deviation from the point of maximum slope is the difference between the position setting for the lower contour and the position of the point of maximum slope for that session, minus the width of the indicator. This difference indicates the extent of the bright band into the region corresponding to constant maximum luminance at the stimulus object. It has been given a negative sign, to emphasize that it represents a deviation from the point of maximum slope in a direction opposite to the upper contour deviation.

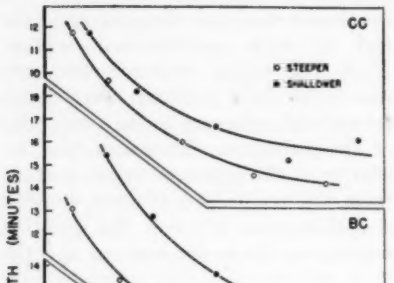

FIG. 5. Width of the apparent bright band as a function of log maximum luminance, for both log luminance gradients. Data are plotted separately for 3 *O*s.

In Figure 6 position data for both gradients are shown separately for each *O*. Zero on the ordinate corresponds to the point of maximum slope of the gradient. Positive values represent positions of contours appearing above the point of maximum slope, in the region of decreasing luminance. Negative values represent positions of contours appearing below this point in the region corresponding to constant luminance at the object.

Details of the results will be considered separately for width of band and for position of contours.

Width of band as a function of luminance.—Each *O* gave data for one or both gradients at five different maximum luminance values. These values cover a range of approximately

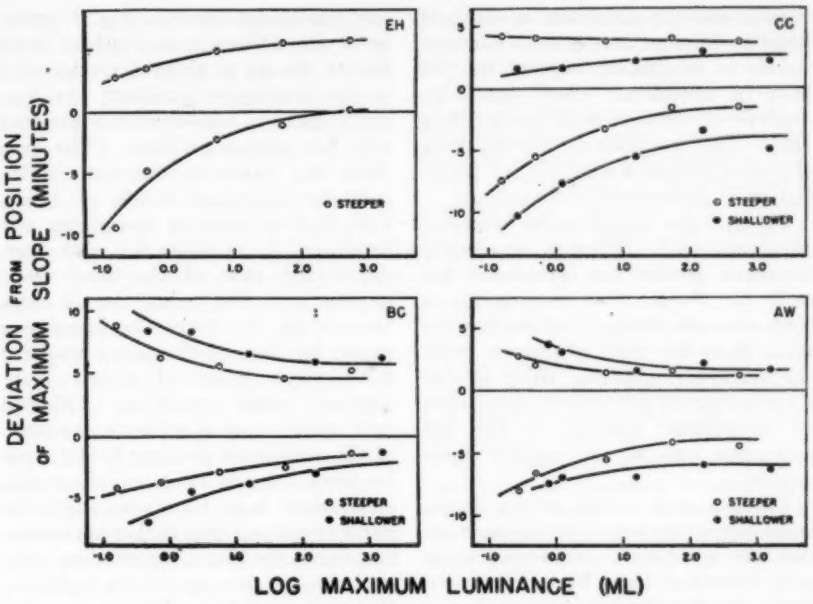

FIG. 6. Position of the apparent bright band with respect to the log luminance gradient, as a function of log maximum luminance. Zero on the ordinate corresponds to the point of maximum slope of the gradient. Contours appearing above this point are shown as positive; contours below this point are shown as negative. Data are plotted separately for four *O*s.

3.5 log ml. At the lower luminances, the values are not the same for all *O*s. Some *O*s were not able to make settings with confidence at the lowest luminances; discriminability of the black indicator line probably acted as a limiting factor in these instances. When *O* was unable to meet the reliability criterion at a low luminance, determinations made at a slightly higher luminance were substituted. In some cases the second lowest luminance level was modified correspondingly in order to space the determinations more evenly.

Each entry in Tables 1 and 2 is based upon settings made within a single experimental session. In some cases, more than one session was devoted to determinations under the same experimental condition. These sessions were equivalent with regard to the experimental conditions, and differences in their results must be attributed to changes in *O*. Since these differences were sometimes large, their results have not been averaged. The data chosen for presentation are those of sessions occurring nearest each other in time. Any other choice, however, would not have reduced the significance of the differences in width.

It is clear from the curves of Fig. 5 that the form of the relation between width of band and luminance is essentially the same for all *O*s, although the absolute magnitude of the effect varies. All show the greatest width at the lowest luminance, with the decrease in width with increasing luminance occurring at a negatively accelerated rate. The change is

regular and the difference in width of band at different maximum luminance values is significant beyond the .01 level of confidence when tested by analysis of variance with ranked data (**10**). This test does not depend upon the assumptions of normality of distribution or homogeneity of variance.

Figure 5 also compares the functions obtained with differing gradients. Separate graphs are presented for each *O*. From these graphs, it is clear that the change in width has the same form for both gradients; with the shallower gradient, there is also some indication that width approaches an asymptote parallel to the log luminance axis at the highest luminances.

Experimental values of log maximum luminance were not the same for the two gradients. However, since both functions have been drawn by plotting width against log maximum luminance, the vertical separation of the curves at any luminance level represents the difference in width of band due to change in the log luminance gradient at that value of maximum luminance. For each *O*, all points for the shallower gradient lie above the curve for the steeper gradient. Since there are no reversals for any of the three *O*s, the data may be regarded as showing a significant increase in width with decreasing steepness of gradient. On the assumption of no difference between the gradients, the probability is $(\frac{1}{2})^5$ or $\frac{1}{32}$ that all five points given by any one *O* for the shallower gradient will lie above those given for the steeper gradient.

These results may also be considered in terms of variations in the luminance gradient, corresponding to those changes described in the log luminance gradient. In these terms, the vertical separation of the curves at any luminance level in Fig. 5 represents the difference in width of band due to change in slope characteristics of the luminance gradient; the narrower band is consistently associated with the steeper gradient. The data show that maximum luminance is also a factor affecting width of band. This is best seen by comparing the height of the ordinate for each open circle with that of the filled circle nearest it to the right. In all cases (except the two lowest log luminance values for CC and the next lowest for AW) these pairs of points were obtained with conditions differing only in the value of α; hence, the slope of the luminance gradient is the same for both. In all 12 cases where this comparison may be made, the filled circle (representing a higher maximum luminance than its corresponding open circle for the same slope) lies higher on the ordinate than the open circle, showing that an increase in maximum luminance when slope is constant increases the width of the apparent bright band.

Position of the band at various luminances.—The data on position of the contours of the bright band are given as deviations from the point of maximum slope of the gradient. Except for one *O*, EH, these deviations are taken from the determination of the position of maximum slope made following the experimental session in which the contour settings were made. For EH, all deviations are from a single determination of the maximum slope position made at the end of the experimental series. As a consequence, the data for EH are less accurately placed with respect to the point of maximum slope than are the data for other *O*s, since the absolute position of that point may be assumed to have changed during the experimental series.

The manner of change of position of the lower contour with increasing luminance is the same for all *O*s, as shown in Fig. 6. The deviation from the position of maximum slope is greatest at low luminances, falling off rapidly at first and then more gradually with increasing luminance. The deviation is greatest with the shallower gradient, with all points for each *O* falling below the curve describing the deviations for the steeper gradient.

Figure 6 shows that *O*s are not in agreement regarding the manner of change of position of the upper contour with increasing luminance. Yet the data for each *O* for each gradient show a clear trend, and the deviations at different maximum luminance values differ significantly beyond the .01 level. Results from multiple sessions also show that *O*s made their adjustments for the upper contour position with great regularity; different values for width obtained in different sessions under the same experimental conditions are in all cases due to a change in judgment of the lower contour position, with the upper contour settings remaining practically identical. It must be concluded that there is a regular change in position of this contour of the band as luminance increases, but that the nature of this change is determined in large part by factors that vary with the individual.

Discussion

The results of this experiment show that there is a regular decrease in apparent width of a bright Mach band as the log luminance gradient giving rise to it is shifted to higher luminances. Two gradients were studied, having different degrees of concavity in the graph describing log luminance with respect to distance. For both gradients, width of the apparent bright band is greatest at the lowest luminance and decreases at a decreasing rate with the shift to higher luminances.

Change from the shallower to the steeper gradient also results in a band of decreased width, when maximum log luminance remains approximately constant. When width is plotted against maximum log luminance for the two gradients of this study, the points obtained with the shallower gradient lie well above those obtained with the steeper gradient.

The experiment shows therefore that width of the apparent band may be altered either by a change in the steepness of the log luminance gradient or by a change in the luminance level at which the gradient occurs.

One characteristic of the stimulus situation is itself altered by both these changes. Consider two points, *a* and *b*, in the stimulus field, such that *a* is any point lying within the uniformly illuminated lower field, and *b* is a point lying at some small arbitrary distance above the point of maximum slope of the luminance gradient. Let the luminances at these two points be I_a and I_b. When the steepness of the log luminance gradient is increased, with I_a constant, the ratio of I_a to I_b is increased. When the log luminance gradient remains unchanged, but the entire gradient is shifted to a higher luminance level, the ratio of I_a to I_b is unchanged. Up to a certain limiting value, such an increase in luminance will cause both I_a and I_b to fall in a luminance range in which smaller ratios are required to detect differences in luminance. The change in width of band with increasing luminance follows a course similar to the course of the change in differential sensitivity of the eye with increasing luminance. It may therefore be stated that conditions which have been shown to decrease the width of Mach bands, i.e., increased luminance and increased slope, also change the relative effectiveness as visual stimuli of luminances at two points, *a* and *b*, chosen to lie to either side of the point of maximum change of luminance.

When the results are examined in terms of the variations in the luminance gradient itself, it is clear that width of band increases both with decreasing slope and with decreasing maximum luminance of the luminance gradient.

The significance of this investigation for certain closely related phenomena should be noted. At the beginning of this report, the stimulus circumstances under which a contour is seen to divide two fields of differing luminance were described. Mach noted that the report of such a contour depends upon a luminance distribution similar in essential respects to luminance distributions giving rise to apparent bands. For when two uniform fields of differing luminance meet at a common border, the discontinuity in the luminance distribution at the physical object is not reproduced at the retina. The distribution at the retina is graded and is sigmoid in form, because of diffraction and aberration in the eye. Gradients having this sigmoid form, but in which change of luminance is more gradual, lead to the report of two Mach bands, one bright and one dark. These bands appear at the "knees" of the function, i.e., in the regions of most rapid change of slope. The upper bend of the function is concave toward the luminance axis and represents a negative change of slope; here a bright band appears. The convex lower bend represents a positive change of slope and gives rise to a dark band. Since the appearance of a band and the establishment of a contour depend on very similar stimulus conditions, it is probable that the processes leading to both phenomena belong to the same class of visual events.

In this context, the present finding that width of bright band decreases with increasing steepness of the luminance gradient[4] may be extended. It may be predicted that band width will continue to decrease with further steepening of the gradient, eventually reaching a width below the acuity threshold. At this point, report of a bright band should give way to the report of a sharp contour between two fields of differing brightness.

Two mathematical treatments of the band phenomenon have been proposed in the literature. Since each of these is, in an important respect, out of accord with data of the present experiment, it is necessary to comment here upon these formulations.

The earlier of these mathematical treatments is that developed by Mach (5). In observing the conditions under which apparent bands occur, he noted the importance of those two characteristics of the luminance gradient upon which width of band has in the present study been shown to depend: the degree of change of slope and the luminance at which a sufficiently large change occurs. His observations led him to conclude that occurrence of a band is a direct function of the degree of change of slope, as measured by the absolute magnitude of the second derivative of luminance with respect to distance $\frac{d^2B}{ds^2}$,[5] and an inverse function of the luminance (B) at the point of this change. The quality of the band (i.e., bright or dark) he described as dependent upon the sign, negative or positive, respectively, of $\frac{d^2B}{ds^2}$.

As a mathematical device for describing the apparent brightness distribution resulting from a given brightness dis-

[4] It should be understood that band width may be described as dependent upon steepness of the gradient to one side of the point of maximum slope only when the gradient to the other side of this point remains unchanged. This was the case in the present experiment, since for all conditions the luminance of the field below the point of abrupt change was constant. If the gradient in this region is also permitted to vary, appearance of the band must be described in relation to some measure of the degree of change of slope.

[5] Mach regarded the use of the second derivative as an approximation, since the mosaic structure of the retina does not allow the assumption of infinitely small gradations in response. He also proposed an equation analogous to the one employing the second derivative, replacing the second derivative term by a term representing radius of curvature of the luminance gradient (6, p. 132).

tribution at the stimulus object, Mach developed an equation based upon Fechner's *Massformel* and employing $\frac{d^2B}{ds^2}$ as a term (**5**, pp. 313–314). The equation is $E = a \log \left[\frac{B}{b} \pm k \frac{\left(\frac{d^2B}{ds^2}\right)^2}{B} \right]$, where E represents apparent brightness, B represents physical intensity, and a and b are Fechner's constants. The sign of the second term within brackets is to be chosen according to the sign of the second derivative. Mach did not regard the equation as a new *Massformel*; it was intended only to "present the facts more briefly and conclusively than can be done in words, without claiming quantitative precision" (**6**, p. 132). He himself observed several difficulties that prevent the equation from achieving such quantitative precision. The equation cannot be used when $\frac{d^2B}{ds^2}$ is infinite; but when the luminance gradient must be described as two distinct, intersecting functions of distance, $\frac{d^2B}{ds^2}$ is always infinite at the point of intersection. Moreover, when these two functions are both of the form $B = Ks$, the formula can predict an increase in apparent brightness only for a single point; whereas, in fact, apparent brightness is increased within a band of finite width.

These objections to the equation are removed by entering in it the physical brightness values at the retina rather than those at the stimulus object. Since the retinal distribution of luminance will always show a rounded gradient, $\frac{d^2B}{ds^2}$ can be evaluated in the cases referred to above. There is, however, a greater inadequacy in the equation which Mach did not indicate. Mach did not discuss variations in band width and did not indicate that he had observed them. As formulated, the equation predicts no change in width of the apparent band as illumination is increased proportionally. This may best be seen by examining the term in the equation which introduces the increment (or decrement) in apparent brightness, the ratio $\frac{\left(\frac{d^2B}{ds^2}\right)^2}{B}$. This term remains constant in value under the changes in the luminance gradient described in this experiment as a shift in position of the log luminance gradient on the log luminance axis. But the present experiment shows that such a shift changes apparent width of the bright band. Although the second derivative may in a new equation prove to be a useful term to describe the apparent brightness distribution, Mach's equation cannot itself describe the band phenomenon adequately.[6]

Fry has suggested (**1**) a mathematical account of the band phenomenon, starting from certain assumptions of Hering's (**3**) regarding the interaction process responsible for the phenomenon. He assumes that every visual element depresses activity in surrounding elements to an extent varying directly with the intensity of its own activity and inversely with its distance from the element undergoing depression. Elements near the edge of a uniformly illuminated area therefore display a higher activity level than elements receiving the same amount of physical stimulation but which lie near the center of the area. The apparent bright band in the vicinity of a sharp decrease in luminance is, on this hypothesis, attributable to the lesser amount of depression exercised on elements in this area by neighboring elements lying in the region of decreased luminance.

As Fry has developed it, this hypothesis regarding interaction predicts only peaks of higher activity, and does not predict bands. Fry assumes a second-order cortical mechanism, which acts upon the distribution resulting from interaction and which has the effect of

[6] A later form of the equation (**8**, pp. 15–18), based on assumptions regarding interaction in the visual system, is subject to the same criticism in the light of present experimental results.

smoothing small differences in activity level. It is not clear that this mechanism can account for changes in band width.

The results of the present experiment give clear information concerning the extent of the bright band into the region corresponding to constant maximum luminance at the object, i.e., lying below the point of maximum slope of the gradient. It has been shown that this extent decreases when the log luminance gradient is shifted to higher luminances. A development of the hypothesis of lessened depression which is capable of predicting bands of finite width should be able to account for this change of extent, by suitable choice of a function relating amount of depression to intensity of activity at the depressing element. However, the data also indicate that the extent of the band into this area decreases with increasing steepness of gradient. It is at this point that the data are most out of accord with the proposed explanation. The hypothesis developed by Fry must predict either no change at all with increasing steepness of gradient, or an increase in extent with increasing steepness.

Summary

1. When the luminance of a stimulus field is graded along one dimension in such a way as to show at some point a sufficiently sharp but continuous decrease in luminance, a bright band is seen in the vicinity of this decrease. The present experiment was designed to investigate this band phenomenon quantitatively. A square stimulus field was used; luminance was constant in the lower half of the field and decreased at a decreasing rate from the center to the top of the field.

2. Measurements of width of the apparent bright band were made at five different values for the logarithm of maximum luminance of the field, over a range of approximately 3.5 log ml. These measurements were made for each of two gradients describing different rates of decrease of the logarithm of luminance within the upper half of the field.

3. When the rate of decrease of log luminance is held constant, and the logarithm of maximum luminance is increased, width of the apparent band decreases according to a negatively accelerated function. When the rate of decrease of log luminance is increased and the logarithm of maximum luminance is held constant, width of the apparent band decreases.

4. The conditions giving rise to report of a contour dividing two fields of differing luminance may be described as a limiting case of that class of luminance gradients which, when the change of luminance is less rapid, gives rise to the report of Mach bands.

5. Two mathematical accounts of the band phenomenon, one developed by Mach and the other by Fry, fail to describe changes in width of band under the conditions giving rise to such changes in the present experiment.

References

1. Fry, G. A. Mechanisms subserving simultaneous brightness contrast. *Amer. J. Optom.*, 1948, **25**, 162–178.
2. Fry, G. A., & Bartley, S. H. The effect of one border in the visual field upon the threshold of another. *Amer. J. Physiol.*, 1935, **112**, 414–421.
3. Hering, E. *Grundzüge der Lehre vom Lichtsinn.* Berlin: Springer, 1920.
4. McDougall, W. Intensification of visual sensation by smoothly graded contrast. *Proc. Physiol. Soc.*, 1903, **1**, 19–21.
5. Mach, E. Ueber die Wirkung der räumlichen Vertheilung des Lichtreizes auf die Netzhaut, I. *Sitz. Wien. Akad. Wiss.*, 1865, **52/2**, 303–322.
6. Mach, E. Ueber die Wirkung der räumlichen Vertheilung des Lichtreizes auf die Netzhaut, II. *Sitz. Wien. Akad. Wiss.*, 1866, **54/2**, 131–144.
7. Mach, E. Ueber die Wirkung der räumlichen Vertheilung des Lichtreizes auf die Netzhaut, III. *Sitz. Wien. Akad. Wiss.*, 1866, **54/2**, 393–409.
8. Mach, E. Ueber die Wirkung der räumlichen Vertheilung des Lichtreizes auf die Netzhaut, IV. *Sitz. Wien. Akad. Wiss.*, 1868, **57/2**, 11–19.
9. Mach, E. Ueber den Einfluss räumlich und zeitlich variierender Lichtreize auf die Gesichtswahrnehmung. *Sitz. Wien. Akad. Wiss.*, 1906, **115/2A**, 633–648.
10. Moses, L. E. Nonparametric statistics for psychological research. *Psychol. Bull.*, 1952, **49**, 122–143.
11. Thouless, R. H. Some observations on contrast effects in graded discs. *Brit. J. Psychol.*, 1922–3, **13**, 301–307.

(Received May 5, 1954)